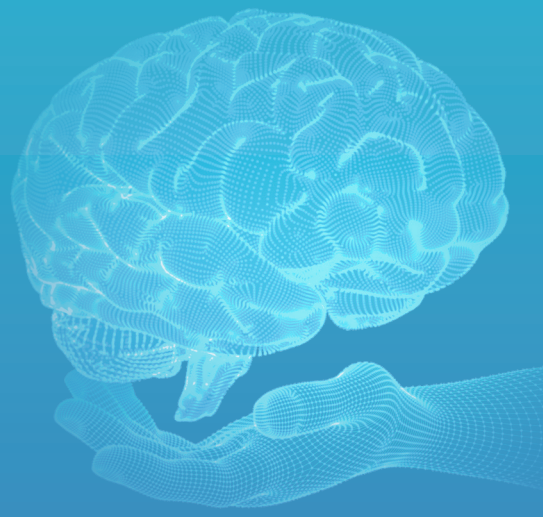
E-ISSN: 3023-784X

Advanced Radiology and Imaging



VOLUME 3 / ISSUE 1
APRIL
2026

galenos Publishing House
ARI
advradiology.org

advradiology.org

Advanced Radiology and Imaging

VOLUME 3 / ISSUE 1
APRIL 2026

EDITORIAL BOARD

VOLUME 3 / ISSUE 1
APRIL
2026

Advanced Radiology and Imaging

advradiology.org

Please refer to the journal's webpage (**https://advradiology.org/**) for "**Journal Policy**" and "**Instructions to Authors**".

The editorial and publication process of the Advanced Radiology and Imaging are shaped in accordance with the guidelines of the **ICMJE, WAME, CSE, COPE, EASE,** and **NISO**. The journal is in conformity with the Principles of Transparency and Best Practice in Scholarly Publishing.

Advanced Radiology and Imaging is indexed in **Türkiye Citation Index, IdealOnline, Zenodo, Scilit,** and **Index of Academic Documents.**

The journal is published online.

Owner: Galenos Publishing House

Responsible Manager: Sonay Aydın

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Türkiye
Phone: +90 (530) 177 30 97 / +90 (539) 307 32 03
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr
Publisher Certificate Number: 14521

Publication Date: April 2026
E-ISSN: 3023-784X
International scientific journal published quarterly.

advradiology.org

Advanced Radiology and Imaging

VOLUME 3 / ISSUE 1
APRIL 2026

CONTENTS

Research Articles

RESEARCH ARTICLE

DOI: 10.4274/AdvRadiolImaging.galenos.2026.79188
Adv Radiol Imaging 2026;3(1):1-7

Microparticle Suspension Intravenous Applications to Enhance Diagnostic Color Doppler Ultrasonography Signal Intensity in the Differential Diagnosis of Liver Tumors

Türkhun Çetin

Erzincan Binali Yıldırım University Faculty of Medicine, Department of Radiology, Erzincan, Türkiye

Abstract

Objectives: To evaluate the effectiveness of color and power Doppler ultrasonography enhanced with the intravenous microparticle contrast agent Levovist in increasing doppler signal intensity and improving differentiation between malignant and benign liver tumors.

Methods: Between December 1999 and July 2000, 21 patients who were admitted to the Gastroenterology and Gastroenterological Surgery Clinics of the Türkiye High Specialization Hospital were included in the study. Patients referred to the Department of Radiology who underwent thick-needle tru-cut biopsy or pathological examination of the postoperative mass were included in the study. Among these 21 patients (12 males and 9 females; age range 29-77 years; mean age 51.4 years), 18 had hepatocellular carcinoma and 3 had cholangiocarcinoma. A total of 19 patients (12 males, 7 females; age 21-71 years; mean 44.9 years) with 9 hemangiomas, 5 liver metastases, 4 hydatid cysts, and 1 liver abscess were selected as the control group.

Results: During the examination, after the Doppler assessment was completed and the ultrasound contrast agent was administered, the values obtained were compared with previous findings. Similar to the literature, the color flow appearances of the masses were divided into four groups: (a) basket pattern (a view where vessels tightly surround the mass and may branch into it), (b) tumor vascularization (vascularization within the mass), (c) dot pattern (a dot-like appearance within the mass), and (d) tortuous pattern (tortuous vascularization surrounding the mass externally) (sinuous vascularization surrounding the mass externally).

Conclusion: As a non-invasive method, contrast-enhanced color Doppler ultrasonography provides important information for the differential diagnosis in selected cases. We believe that, as color and power Doppler devices are further developed to allow examination of deep-seated and small intraparenchymal tumoral hepatic lesions and to reveal blood flow velocity in very small vascular structures, examinations using contrast agents to increase Doppler signal intensity will provide diagnostically important information.

Keywords: Doppler, liver, ultrasonography, intratumoral vascularization, contrast-enhanced doppler ultrasonography, hepatocellular carcinoma

Introduction

Grey-scale ultrasonography (US), color Doppler ultrasonography (CDUS), and power Doppler ultrasonography are the preferred noninvasive methods in the evaluation of liver masses. However, these methods may not always characterize the lesion. In such cases, triphasic dynamic computed tomography (CT) or magnetic resonance imaging (MRI) is performed for further evaluation. However, the lesion can also be characterized by dynamic contrast US and CDUS with echocontrast agents.

There are two important phases during examination of the liver with echocontrast agents[1,2]: the early vascular phase (arterial and portal phases) and the late parenchymal phase.

In CDUS examinations performed with IV contrast agents, assessment of the accuracy of quantitative perfusion parameters for the effective differential diagnosis of malignant and benign liver lesions is among the most important diagnostic criteria. Some echocontrast agents are retained in cells of the reticuloendothelial system during the late parenchymal phase (Levovist, Sonazoid, among others), and this retention lasts approximately 20 minutes.

Benign lesions show contrast enhancement in both the early vascular phase and the late parenchymal phase. However, a low perfusion defect can be observed in the late phase in some benign lesions. Malignant lesions and metastases usually show contrast enhancement in the early vascular phase, whereas in the late phase they lack enhancement and appear as contrast defects. However, persistent perfusion may be

Cite this article as: Çetin T. Microparticle suspension intravenous applications to enhance diagnostic color doppler ultrasonography signal intensity in the differential diagnosis of liver tumors. Adv Radiol Imaging. 2026;3(1):1-7

Address for Correspondence: Asst. Prof. Türkhun Çetin, Erzincan Binali Yıldırım University Faculty of Medicine, Department of Radiology, Erzincan, Türkiye
E-mail: turkhuncetinmd@gmail.com ORCID ID: orcid.org/0000-0003-0209-4218
Received: 12.01.2026 Accepted: 23.02.2026 Epub: 05.03.2026 Published: 30.04.2026

observed infrequently in the late phase of malignant lesions. Peripheral contrast uptake may also be present in late-phase examinations of metastases.

In this context, ultrasound contrast agents contribute to the characterization of liver masses and offer an alternative to CT and MRI. Side effects of echocontrast agents are very rare, transient, and mild.[2,3] They do not have specific cerebral, hepatic, or renal toxicities. Therefore, dynamic contrast-enhanced US and Doppler examinations may be preferred in patients with hepatic and renal insufficiency, and in cases where MRI and CT cannot be performed.

SHU 508 A (Levovist), the echocontrast agent used in our study, contains specific microbubbles, galactose, and palmitic acid molecules. Its side effects are nonspecific symptoms, including sensations of heat or cold, injection-site pain, and taste changes. The only contraindication is galactosemia. Its safety profile is considered acceptable.

In this study, the effectiveness of contrast-enhanced and Doppler US in the differential diagnosis of hepatic neoplastic masses and hepatic metastases was investigated, and digital subtraction angiography (DSA) results were compared with vascularization findings.

This research thesis was conducted in the Department of CDUS and in the DSA Unit of the Radiology Clinic at the Türkiye High Specialization Hospital (THIY). The study included 21 patients with primary liver tumors, diagnosed by clinical and pathological examinations, and a control group of 19 patients with intraparenchymal space-occupying lesions in the liver.

In the patient and control groups, color and power Doppler examinations were performed before and after IV administration of Levovist (SHU 508 A), a contrast agent with a molecular structure based on galactose and palmitic acid. Peripheral and intratumoral vascular flow patterns, waveform characteristics, arterial flow velocity, and resistive index (RI) of the mass lesion were examined and recorded.[2-4]

Ultrasound with intravenous contrast agents, including B-mode and Doppler modes, can be performed during the same examination session, providing additional diagnostic information and significantly improving diagnostic performance compared with dynamic CT and MRI with respect to vascularization features and specific contrast-dynamic characteristics.[3,5]

Methods

All procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards (approval no.: 80576354-050-99/169, date: 29.05.2019). The study was approved by the TYIH Chief Medical Officer Training and Planning Commission and conducted in accordance with the Declaration of Helsinki. Patient consent was obtained.

Between December 1999 and July 2000, 21 patients admitted to the Gastroenterology and Gastroenterological Surgery Clinics of the TYIH were included in the study. Patients who were referred to the Radiology Department and who underwent thick-needle tru-cut core biopsy or who underwent pathological examination of the postoperative mass were selected for inclusion in the study. Of these 21 patients (12 males, 9 females; age range 29-77 years; mean age 51.4 years), 18 had hepatocellular carcinoma (HCC) and 3 had cholangiocarcinoma. A total of 19 patients (12 males, 7 females; age range 21–71 years; mean age 44.9 years) diagnosed with hemangiomas (n=9), liver metastases (n=5), hydatid cysts (n=4), and liver abscess (n=1) were selected as the control group.

All patients included in the study were examined using a LOGIQ 700 MR Doppler US device (GE Medical Systems) with convex and sector probes. Hepatic mass lesions occupying the intraparenchymal space were initially evaluated by B-mode US for size, contour, borders, and echo pattern. Subsequently, a color Doppler examination was performed. In masses showing tumoral vascularization, peripheral and intratumoral color Doppler flow patterns and appearances; waveform; feeding artery; arterial flow velocity; RI; hepatic artery flow velocity; and hepatic perfusion index were evaluated. The US device was adjusted to detect low flows by lowering the color filter and reducing the PRF and wall filter.

During the examination, after completion of the Doppler evaluation, Levovist (galactose palmitic acid-containing granules, 999 mg) was administered intravenously at a concentration of 400 mg/mL in a total volume of 20 mL, at a constant injection rate of 2 mL/sec. This provided a homogeneous increase in the Doppler signal lasting an average of 90 seconds, with maximum effect at 30-60 seconds. Subsequently, the values obtained after administration of the ultrasound contrast agent were compared with the previous findings.[1,4]

The Tanaka et al.[5] classification served as the basis for the study, and color-flow appearances within the masses were divided into four groups. According to this classification,[3,5] the following patterns were investigated: (a) basket pattern (an appearance that tightly surrounds the mass and may give rise to branches into the mass);(b)tumoral vascularization(presence of vascularization within the mass); (c) spot pattern (a punctate appearance within the mass); and (d) detouring pattern (a vascular pattern surrounding the mass in a detouring manner).

Spectra were obtained from colored flow images at Doppler angles between 30° and 60°, and flow velocities and characteristics were determined. Blood flow waveforms were classified into two groups based on their spectral patterns: pulsatile and continuous (Table 1). The mean and median flow velocities within the masses were determined.[5,6]

No significant side effects or adverse events were observed following intravenous administration of Levovist (400 mg/mL), 20 mL, at an injection rate of 2 mL/sec. Only one patient , a 71-year-old, experienced pain and increased temperature at the injection site, followed by nausea; these symptoms resolved within 10 minutes without intervention.[7,8]

Findings on selective celiac and hepatic artery DSA included arterial dilatation and elongation, coarse, dense neovascularization, vascular pooling, and A.V. shunts. Dense neovascularization, staining, contrast enhancement, and tense (stretched) abnormal vascularization may

Table 1. All flow grades and patterns on color doppler ultrasonography

Doppler flow degree	CDUS finding (sign)
0	No current
1	Normal organ blood flow
2	Increased blood flow
3	Patterns of excessive blood supply

CDUS: Color doppler ultrasonography

be observed around areas of necrosis. The lesion may also appear totally hypovascular. DSA alone is sufficient for evaluating vascular invasion and preoperative mapping of the vasculature (e.g., portal vein invasion). CDUS with Levovist can evaluate the feeding artery before chemoembolization. Cholangiocarcinoma appears angiographically similar to anaplastic HCC. It is a highly hypovascular tumor that causes encirclement, occlusion, and displacement of the primary intrahepatic branches of the hepatic artery and the portal venous system.[8,9] Neovascularization is present in approximately 50% of patients, while tumor contrast staining is absent or very weak. The most common locations are the proximal and distal branches of the left and right common hepatic arteries.

Statistical Analysis

The data were analyzed using the Statistical Package for the Social Sciences (SPSS) for Windows, version 20 (SPSS Inc., Chicago, IL, USA). The normality of data distribution was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Normally distributed numerical variables were presented as mean ± standard deviation, whereas non-normally distributed variables were presented as minimum-maximum values.

The diagnosis of liver lesions in the control group was made clinically and histopathologically, and in some cases was confirmed by laboratory tests and US. For statistical analysis, the Mann-Whitney U test was used because the number of patients in the control group was small. The Student's t-test was used for comparisons between two groups. Friedman's test and Tukey-Kramer multiple-comparison tests were used to compare cases with angiographic findings and those with Power Doppler findings.

Results

The study included 21 patients with a mean age of 51.4 years, of whom 12 were female and 9 were male (Table 2).

When contrast uptake in vascular phases, perfusion properties in late phases, and CDUS and power Doppler characteristics of all hepatic lesions were evaluated, 26 of the 39 cases with color Doppler flow appearance were malignant (96% of all malignant masses), and 13 were benign (13.6% of all benign masses).

The mean maximum blood flow velocities ranged from 7 to 158 cm/s in malignant lesions and from 6 to 39 cm/s in benign lesions. Statistical comparison of mean blood flow velocities revealed statistically significant differences between benign and malignant masses (t: 8.15,

Table 2. CDUS acuity pattern distribution of 21 patients with primary liver tumours and 19 patients in the control group

Cases	Basket pattern	Tumoral Vascularization	Detouring pattern	Spot pattern
HCC (18 patients)	14	4		
Cholangiocarcinoma (3 patients)		2	1	
Liver metastasis (5 patients)		3	2	
Hepatic haemangioma (9 patients)		0	1	7
Liver abscess (1 patient)			1	
Hepatic cyst (4 patients)				

CDUS: Color Doppler ultrasonography

Table 3. Rates of intratumoral vascularisation discovered in the intrahepatic parenchymal lesions examined (number of cases and percentage of vascularization revealed)

Type of intrahepatic nodular lesion	Total lesion	Angiography (WHO)	Conventional colour doppler usg	iv. contrast agent After application (400 mg/mL) Colour doppler usg signal intensity increase
HCC	18 Patients	14* (77)	9 (50)	13* (72)
< 20 mm. HCC	(4)	(2)	(0)	(1)
20-30 mm. HCC	(7)	(5)	(4)	(5)
≥30 mrn. HCC	(7)	(7)	(5)	(7)
Hepatic hemangioma	9 Patients	9 (100)	4 (44.5)	8 (89)
Hepatic metastasis	5 Patients	3 (60)	2 (40)	3 (60)
Cholangiocarcinoma	3 Patients	1 (33)	1 (33)	2 (66)
Total	**35**	**27 (77)**	**16* (46)**	**26* (74)**

WHO: World Health Organization

p<0.001) and between hepatoma and hemangioma (t: 3.38, p<0.001). A significant difference (p<0.05) between HCC and metastatic cases was also observed.

In 13 cases, significant differences in vascularization and Doppler patterns were observed after IV administration of Levovist (galactose + palmitic acid). In these cases, the pre-administration flow grade was 1 (normal organ blood flow), and after administration, a flow grade of 2 (increased blood flow) or 3 (excessive blood flow) was detected (Table 3).

Of the 18 patients diagnosed with HCC, 14 exhibited a basket pattern (77%), and the remaining 4 showed tumoral neovascularization (23%). Tumor sizes measured by US ranged from 23-167 mm (mean, 56 mm). Maximum and minimum flow velocities ranged from 61-158 cm/s (mean 96.35 cm/s) and 0-53 cm/s (mean 32.4 cm/s), respectively. Flow grade ranged from 1 to 3 (mean=2.2). Peak systolic maximum ranged from 1.17-5.2 kHz (mean 3.8 kHz), and peak systolic minimum ranged from 0-3.1 kHz (mean 0.99 kHz).

Discussion

Applications of CDUS were developed by Namekawa et al. and used to investigate major cardiac vessels and intracardiac blood flow.[10] In 1985, Sukigara et al.[11] investigated 11 patients with HCC using CDUS.[12] Although angiography and US provide important diagnostic information for liver tumors, CDUS can provide additional information not revealed by either method.

In CDUS, the color of the blood flow within the mass depends on the flow rate in the vessel and not on vessel diameter or blood volume. Therefore, even thin vessels whose blood flow velocities are not visible on B-mode imaging can be visualized with CDUS, allowing spectral analysis and acquisition of quantitative data.

Our study was based on the Tanaka et al.[5] classification.[9,10] According to this classification, four signs can be identified in liver tumors using CDUS; the basket pattern and tumoral vascularization are characteristic findings of HCC. These two signs were not observed in hemangiomas or metastatic liver tumors. In our study, the sign of tumoral vascularization with a basket-like pattern was found only in cases of HCC, and this result is consistent with the findings of Tanaka et al.[5,9,13,14]

Ueno et al.[9] reported that they could not detect CDUS blood flow in HCC cases smaller than 2 cm,whereas Tanaka et al.[5] detected CDUS blood flow in five cases smaller than 2 cm, attributing this to the greater technological capability and sensitivity of the CDUS device they used. Taylor et al.[10] reported high flow velocity in cases of HCC and attributed this finding to wide pressure gradients across arteriovenous shunts within the tumor. Tanaka et al.[5] reported that the maximum flow velocity in tumor arteries ranged from 70 to 90 cm/s. In our study, although a wider range (61-158 cm/s; mean 96.35 cm/s)was observed, the mean values were comparable.

The direction of flow within the tumor was pulsatile in 13 cases (72%) and continuous in 5 cases (28%). Tanaka et al.[5] reported pulsatile flow in 80% of cases and continuous flow in 20%, indicating that the findings of our study are similar to theirs.[8,11]

The absence of a basket pattern and tumoral vascularization in hemangiomas on CDUS significantly differentiated hemangiomas from HCC; therefore, CDUS can be used in the differential diagnosis between hemangiomas and HCC. In our hemangioma patients, a spot pattern was observed in seven cases and a detouring pattern in one case. In one hemangioma case, no Doppler signal was detected.[7,11]

Among five control-group cases with liver metastases, tumoral neovascularization was observed in three and a detouring pattern in two. Maximum blood flow velocities ranged from 7 to 125 cm/s (mean, 59.12 cm/s). Continuous flow was observed in six hemangioma cases. None of the benign lesions other than hemangiomas (abscesses, hepatic cysts, hydatid cysts, and others) showed Doppler shifts or pathological signals.

Overall, among the 39 cases showing color Doppler flow, 26 were malignant (representing 96% of all malignant masses), and 13 were benign (representing 13.6% of all benign masses). Mean maximum blood flow velocities ranged from 7-158 cm/s in malignant lesions and 6-39 cm/s in benign lesions. Statistical comparison of mean blood flow velocities revealed highly significant differences between benign and malignant masses (t: 8.15, p<0.001) and between hepatomas and hemangiomas (t: 3.38, p<0.001). There was also a significant difference (p<0.05) between hepatomas and metastatic cases.

In the study, 21 patients with primary liver tumors (18 with HCC and 3 with cholangiocarcinoma) received IV Levovist (400 mg/mL, 20 mL) after color Doppler examination; their CDUS evaluations were then repeated. Flow patterns and signal intensity values obtained after IV administration of microparticle suspension (Levovist; galactose + palmitic acid) were compared separately with those obtained by conventional CDUS and Power Doppler US examinations.[6,8,15]

In 13 cases, significant differences were observed in lesion vascularization, Doppler findings, and flow velocities after IV Levovist administration. In these cases, conventional CDUS showed a flow grade of 1 (normal organ blood flow) before administration and grades 2 (increased blood flow) or 3 (excessive blood flow) after administration.

Across all cases, the application of Levovist during conventional CDUS resulted in varying degrees of clarification of the specific and pathognomonic flow patterns of tumor vascularization, particularly at the lesion periphery and within the lesion, manifesting as increased Doppler signal. Compared with the Power Doppler feature, only five cases demonstrated a significant signal increase or other findings that would result in a diagnostic change.[6,8,15]

According to the results of these Doppler examinations, although intravenous administration of microparticle suspensions such as Levovist did not produce a statistically significant difference in advanced CDUS devices with Power Doppler features, evaluation of the five responsive cases suggests that it is more effective, particularly for diagnosing primary liver tumors located deep within the liver and smaller than 3 cm. However, with conventional CDUS devices, IV Levovist increases Doppler signal intensity to varying degrees (p<0.05).

According to our study results and the Tanaka et al.[5] classification, the possibility of HCC should be considered in cases showing a basket pattern and tumoral vascularization, whereas hepatic hemangioma should be considered in cases showing a spot pattern. Color-flow appearance and neovascularization are observed on CDUS in 90-100% of malignant masses. However, the absence of a Doppler signal does not rule out malignancy. When maximum flow velocities are compared statistically, a significant difference is observed between malignant and benign liver masses.[9,11,15]

In conclusion, CDUS examinations provide important information for the differential diagnosis of liver tumors, based on blood-flow velocities and flow-pattern characteristics. However, CDUS may be inadequate for imaging blood flow when tumors are located deep within the liver

parenchyma. Therefore, there is a need for CDUS technologies capable of evaluating smaller and deeper tumors, detecting blood flow velocity in very thin vessels, and employing echocontrast agents to significantly increase Doppler signal intensity.

selective celiac and hepatic artery DSA examinations of patients with primary liver tumors (18 HCC and 3 cholangiocarcinoma) revealed arterial enlargement and elongation, coarse neovascularization, vascular ponding, and intratumoral AV shunts in HCC cases. In our study, 14 of the 18 HCC patients exhibited intense neovascularization, pathological staining, contrast enhancement, and stretched abnormal neovascularization around necrotic areas; total hypovascularity was observed in four cases. When these findings were compared with Color- and Power-CDUS results obtained after IV administration of Levovist 400 mg/mL (20 mL), a significant correlation with digital selective angiography findings was observed.[5,8]

Evaluation of vascularity in primary liver lesions by selective celiac and hepatic artery DSA examinations demonstrated a strong correlation and agreement with color and power CDUS findings obtained after administration of Levovist. Fujimoto et al.[12], demonstrated that pulsatile arterial flow and Doppler flow-pattern findings were significantly correlated with arterial angiography results in HCC patients who received IV SH-U-508A contrast agents at three different concentrations (200, 300, and 400 mg/mL).

In that study, administration of the contrast agent SH-U-508A at 400 mg/mL IV increased the prominence of intratumoral arterial flow-pattern features by 83% in HCC patients. Particularly in HCC lesions smaller than 3 cm, the contrast agent significantly enhanced intratumoral flow signals, increasing their intensity.[8,13,16]

In the present study, a basket pattern was observed in the majority of HCC cases, which is consistent with previous reports. Similar vascular patterns and increased peak systolic velocities in malignant liver lesions have been reported in earlier studies. These findings support the concept that malignant tumors tend to exhibit arterial vascularization that is more chaotic and more prominent than that of benign lesions. The significant increase in Doppler signal intensity after contrast administration further emphasizes the role of contrast-enhanced Doppler US in differentiating malignant from benign hepatic tumors.

Digital selective angiography did not reveal intratumoral arterial flow in 15 cases of hypovascular HCC (73%). Focal nodular lesions, particularly those smaller than 2 cm, cannot be reliably distinguished on angiography because of insufficient contrast enhancement. Metastatic lesions can be visualized as hypovascular areas in the parenchymal phase, although intratumoral vascularization may not be detected, they can appear surrounded by relatively enlarged or tortuous arterial structures.

In all cases, intrahepatic blood supply and Doppler signal intensity were enhanced in conventional color Doppler sonographic examinations after IV administration of the SH-U-508A (Levovist®) contrast agent (400 mg/mL).[6,15,17] Color and Power Doppler sonography are now widely used in the differential diagnosis and evaluation of the hemodynamic characteristics of hepatic tumors. As non-invasive methods, they provide important diagnostic information for HCC and other intrahepatic parenchymal lesions and enable the detection of arterial flow characteristics in hypervascular lesions. Wilson SR et al.[17] reported diagnostic accuracies of 85–92% when qualitative CDUS values obtained after IV contrast administration were evaluated using an algorithm-based approach.[16,17]

Beyer et al.[18] differentiated focal benign from malignant hepatic mass lesions by quantifying Doppler blood flow, blood volume, and peak enhancement, reporting receiver operating characteristic (ROC) values of 0.97, 0.96, 0.98, and 0.76 (n = 20).

With further technological improvements in color and power Doppler systems, these Doppler methods may provide information comparable to CTA and digital selective angiography. However, the main limitations of CDUS include the influence of deeply located lesions and background vascular structures on flow signals, which may hinder adequate examination.[9,16,18]

In 1997, Tano et al.[9], investigated the enhancement of Doppler signal intensity using CDUS with a galactose-based contrast agent to differentiate HCCs smaller than 20 mm from hemangiomas and focal fatty areas. Intratumoral color Doppler images were evaluated before and after IV administration of the contrast agent SH-U-508A (Levovist®) (400 mg/mL).[6,13,18] After injection, intratumoral and peripheral vascularization were enhanced in all HCC cases and in two hemangioma cases. They reported that enhancement of Doppler signal intensity using the SH-U-508A contrast agent may be useful in the differential diagnosis of small hyperechogenic HCC lesions from other hyperechogenic parenchymal lesions (mainly hemangiomas and focal fatty areas).[6,16,17]

Recent single-center investigations, including a study by Gatos et al.[19] that proposed a classification algorithm for intrahepatic lesion contrast patterns using computer-assisted evaluations (n=52), have reported diagnostic accuracy rates for Doppler US reaching up to 90.3%.

Furthermore, a clinical study conducted by Goertz et al.[20] (n=33) demonstrated that quantified contrast intensity parameters obtained during the late phases were significantly lower in malignant focal liver lesions than in their benign counterparts.

This phenomenon is attributed to the rapid washout of Doppler contrast from the lesion, which is currently regarded as a well-established qualitative criterion for malignancy.

In a subsequent study (n=148), Wildner et al.[21] quantified late-phase Doppler contrast enhancement patterns across various tumor subgroups—including hepatic focal nodular hyperplasia, HCC, cholangiocellular carcinoma, and hepatic metastases of breast, pancreatic, and colon cancers—revealing statistically significant differences among these cohorts.

Both maximum and minimum flow velocities were significantly higher in malignant lesions than in benign lesions ($p<0.05$). Pulsatile flow patterns were more frequently observed in malignant tumors. Following Levovist administration, Doppler signal intensity increased in benign and malignant lesions, although the increase was statistically significant only in malignant lesions. These findings indicate that contrast-enhanced Doppler US improves the detection and characterization of malignant vascular patterns.

Measurements of flow were obtained at locations that were both central and peripheral in 9 cases, central only in 5 cases, and both inside and outside the tumor in 2 cases.

The enhancement of Doppler signal intensity following intravenous contrast administration may provide additional diagnostic value in cases where baseline Doppler findings are inconclusive. Therefore, contrast-enhanced Doppler US may serve as a useful adjunct in the

differential diagnosis of liver tumors, particularly for distinguishing malignant lesions based on vascular characteristics.

Study Limitations

Statistical power could be improved with a larger sample size. Doppler ultrasonographic RI and PI values could also be included. Hepatocyte-specific contrast agents such as Gd-EOB-DTPA could improve correlation between contrast-enhanced dynamic liver MRI and multislice CT imaging.

This study reflects a single-center clinical experience; multicenter studies could broaden the scope.[15,21,22]

Future research involving larger cohorts should evaluate perfusion quantification parameters using intravenous contrast-enhanced Doppler imaging to identify the most effective parameters for differentiating malignant from benign liver lesions.

Additionally, it is imperative to establish statistically significant correlations with standard radiological and diagnostic references, such as histopathological examinations and dynamic liver MRI.

The advent of novel intravenous contrast agents for Doppler US and innovative particulate chemical solutions holds substantial promise to improve diagnostic accuracy in differentiating malignant from benign parenchymal lesions.

Conclusions

With conventional color Doppler sonography, arterial flow and Doppler patterns can be visualized more clearly when contrast agents are used at higher concentrations, and findings can be correlated with selective digital angiography.

As new contrast agents and advanced technologies for IV contrast-enhanced CDUS are developed, and as dynamic CT, MRI, and invasive DSA advance, the differential diagnosis of intrahepatic mass lesions will become clearer, and treatment protocols will be determined more accurately. Technological advancements in Doppler US devices are expected to further contribute to this field.

Ethics

Ethics Committee Approval: The study was approved by the Türkiye High Specialisation Hospital Chief Medical Officer Training and Planning Commission. All procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards (approval no.: 80576354-050-99/169, date: 29.05.2019).

Informed Consent: Written formal consent for the publication of scientific data and related information has been obtained from all patients included in this scientific study (i.e., the study participants or, where applicable, their first-degree relatives) and has been recorded in the hospital database.

Footnotes

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Strobel D, Hoefer A, Martus P, Hahn EG, Becker D. Dynamic contrast-enhanced power Doppler sonography improves the differential diagnosis of liver lesions. Int J Colorectal Dis. 2001;16:247-56.
2. Tamsel, İ. The effectiveness of color doppler ultrasonography and echo contrast agents in the differential diagnosis of liver masses. Ege Journal of Medicine. 2020;59:241-50.
3. Yang CC, Chen CH, Yeh YH. Enhanced color flows in hepatic tumors. Hepato-gastroenterology. 2002;49:1506-9.
4. Tranquart F, Bleuzen A, Kissel A. Value of combined conventional and contrast-enhanced sonography in the evaluation of hepatic disorders. Journal de Radiologie. 2004;85:755-62.
5. Tanaka S, Yoshioka F, Kitamra T, Uchimoto R, Niwa K, Miyazawa T. [Contrast enhanced color Doppler sonogram of liver tumors: a color-filled pattern in the late phase]. Nihon Rinsho. 1998;56:961-6. Japanese.
6. Bartolotta TV, Midiri M, Quaia E, et al. Liver haemangiomas undetermined at grey-scale ultrasound: contrast-enhancement patterns with SonoVue and pulse-inversion US. Eur Radiol. 2005;15:685-93.
7. Choi BI, Kim AY, Lee JY, Kim KW, Lee KH, Kim TK, Han JK. Hepatocellular carcinoma: contrast enhancement with Levovist. J Ultrasound Med. 2002;21:77-84.
8. Pei XQ, Liu LZ, Zheng W, et al. Contrast-enhanced ultrasonography of hepatocellular carcinoma: correlation between quantitative parameters and arteries in neoangiogenesis or sinusoidal capillarization. Eur J Radiol. 2012;81:e182-8.
9. Tano S, Ueno N, Tomiyama T, Kimura K. Possibility of differentiating small hyperechoic liver tumours using contrast-enhanced colour doppler ultrasonography: a preliminary study. Clin Radiol. 1997;52:41-5.
10. Taylor KJ, Ramos I, Carter D, Morse SS, Snower D, Fortune K. Correlation of doppler US tumor signals with neovascular morphologic features. Radiology. 1988;166:57-62.
11. Sukigara M, Koga K, Komazaki T, Omoto R. Clinical experience with intraoperative doppler color flow imaging in hepatectomy. J Clin Ultrasound. 1987;15:9-15.
12. Fujimoto M, Moriyasu F, Nishikawa K, Nada T, Okuma M. Color doppler sonography of hepatic tumors with a galactose-based contrast agent: correlation with angiographic findings. AJR Am J Roentgenol. 1994;163:1099-104.
13. Strobel D, Krodel U, Martus P, Hahn EG, Becker D. Clinical evaluation of contrast-enhanced color Doppler sonography in the differential diagnosis of liver tumors. J Clin Ultrasound. 2000;28:1-13.
14. Namekawa K, Kasai C, Tsukamoto M, Koyano A. Realtime bloodflow imaging system utilizing auto-correlation techniques. Ultrasound Med Biol. 1983;Suppl 2:203-8.
15. Chaubal N, Joshi M, Bam A, Chaubal R. Contrast-enhanced ultrasound of focal liver lesions. Semin Roentgenol. 2016;51:334-57.
16. Battaglia V, Cervelli R. Liver investigations: Updating on US technique and contrast-enhanced ultrasound (CEUS). Eur J Radiol. 2017;96:65-73.
17. Wilson SR, Burns PN. An algorithm for the diagnosis of focal liver masses using microbubble contrast-enhanced pulse-inversion sonography. AJR Am J Roentgenol. 2006;186:1401-12.
18. Beyer LP, Pregler B, Wiesinger I, Stroszczynski C, Wiggermann P, Jung EM. Continuous dynamic registration of microvascularization of liver tumors with contrast-enhanced ultrasound. Radiol Res Pract. 2014;2014:347416.
19. Gatos I, Tsantis S, Spiliopoulos S, et al. A new automated quantification algorithm for the detection and evaluation of focal liver lesions with contrast-enhanced ultrasound. Med Phys. 2015;42:3948-59.

20. Goertz RS, Bernatik T, Strobel D, Hahn EG, Haendl T. Software-based quantification of contrast-enhanced ultrasound in focal liver lesions--a feasibility study. Eur J Radiol. 2010;75:e22-6. Wildner D, Schellhaas B, Strack D, et al. Differentiation of malignant liver tumors by software-based perfusion quantification with dynamic contrast-enhanced ultrasound (DCEUS). Clin Hemorheol Microcirc. 2019;71:39-51.

21. Wildner D, Schellhaas B, Strack D, et al. Differentiation of malignant liver tumors by software-based perfusion quantification with dynamic contrast-enhanced ultrasound (DCEUS). Clin Hemorheol Microcirc. 2019;71:39-51.

22. Kong WT, Ji ZB, Wang WP, Cai H, Huang BJ, Ding H. Evaluation of liver metastases using contrast-enhanced ultrasound: enhancement patterns and influencing factors. Gut Liver. 2016;10:283-7.

RESEARCH ARTICLE

DOI: 10.4274/AdvRadiolImaging.galenos.2026.44153
Adv Radiol Imaging 2026;3(1):8-14

Reassessment of CT Images in Choledocholithiasis: Analysis of Missed Findings and Associated Factors

Bahadır Reis, Faruk Enes Yazar

Kastamonu University Faculty of Medicine, Department of Radiology, Kastamonu, Türkiye

Abstract

Objectives: Computed tomography (CT) is frequently the first cross-sectional imaging modality obtained in patients presenting with suspected biliary obstruction, yet its sensitivity for direct detection of common bile duct (CBD) stones is known to be variable. This study aimed to determine the rate of CT-missed findings in patients with magnetic resonance cholangiopancreatography (MRCP)-confirmed choledocholithiasis and to identify stone and protocol characteristics associated with CT nondetection.

Methods: In this single-center retrospective study, 33 adults with MRCP-confirmed choledocholithiasis who had undergone abdominal CT within seven days of MRCP were included. CT images were independently re-evaluated by two blinded radiologists; discordant readings were resolved by consensus. MRCP served as the reference standard. CT sensitivity and its 95% confidence intervals (CI) were calculated using the Wilson method. Intergroup comparisons were performed using the Mann-Whitney U test and Fisher's exact test. Agreement between MRCP and CT measurements of CBD diameter was assessed by Spearman correlation and Bland-Altman analysis.

Results: CT sensitivity was 48.5% (16/33; 95% CI, 32.5-64.8%), corresponding to a missed finding rate of 51.5%. MRCP stone diameter was significantly larger in the CT-detected group (median 9.0 mm vs. 5.0 mm; p=0.001), and CT-measured CBD diameter was likewise larger (median 14.0 mm vs. 10.0 mm; p=0.017). Stones measuring ≤5 mm were missed on CT in 83.3% of cases versus 33.3% for stones larger than 5 mm (odds ratio, 10.00; p=0.010). MRCP and CT CBD diameters were strongly correlated (Spearman r=0.793; p<0.001), with negligible bias on Bland-Altman analysis (+0.18 mm). The Median stone attenuation in CT-detected cases was 78 hounsfield units (HU) (interquartile range, 56-114 HU).

Conclusion: CT failed to identify CBD stones in more than half of patients with MRCP- confirmed choledocholithiasis. Stone diameter ≤5 mm was the strongest predictor of CT non-detection. These findings support a low threshold for MRCP referral when CT does not directly demonstrate a CBD stone in patients with clinically suspected biliary obstruction.

Keywords: Choledocholithiasis, computed tomography, magnetic resonance cholangiopancreatography, sensitivity, missed diagnosis, common bile duct

Introduction

Cholelithiasis affects approximately 10-15% of the adult population in Western countries, and choledocholithiasis the presence of stones within the common bile duct (CBD) develops in 10-20% of patients with gallbladder stones at some point during their disease course.[1,2] Left unrecognised or inadequately managed, CBD stones predispose to serious biliary complications including acute cholangitis, acute biliary pancreatitis, and secondary biliary cirrhosis, all of which are associated with substantial morbidity and, in severe cases, mortality.[3] Timely identification and appropriate triage of patients for further evaluation by magnetic resonance cholangiopancreatography (MRCP), endoscopic ultrasound (EUS), or therapeutic endoscopic retrograde cholangiopancreatography (ERCP) is therefore a clinical priority.[4]

In current clinical practice, transabdominal ultrasonography is the first line imaging modality for suspected biliary pathology; however, its sensitivity for direct CBD stone visualisation is limited, ranging from 22% to 55% in published series, largely owing to acoustic shadowing from overlying bowel gas and operator dependence.[5,6]

Abdominal CT is widely obtained in emergency department patients presenting with non-specific abdominal pain, right upper quadrant discomfort, jaundice, or suspected acute pancreatitis, owing to its broad anatomical coverage and rapid acquisition.[7] While CT can simultaneously evaluate a wide differential diagnosis and detect indirect biliary signs notably CBD dilatation the direct visualisation of a CBD stone on CT is considerably less reliable and is influenced by a complex interplay of stone composition, size, location, and technical imaging factors.[8-10]

Cite this article as: Reis B, Yazar FE. Reassessment of CT images in choledocholithiasis: analysis of missed findings and associated factors. Adv Radiol Imaging. 2026;3(1):8-14

Address for Correspondence: Bahadır Reis MD, Kastamonu University Faculty of Medicine, Department of Radiology, Kastamonu, Türkiye
E-mail: bahadirreis@hotmail.com ORCID ID: orcid.org/0009-0003-7260-6917
Received: 19.02.2026 Accepted: 24.04.2026 Published: 30.04.2026

The reported sensitivity of CT for choledocholithiasis varies widely across the literature, from approximately 40% to 80% depending on the study cohort, CT protocol, and reference standard used.[8-11] Isoattenuating stones, stones smaller than 5 mm, and stones situated in the distal CBD are particularly prone to being overlooked.[8,10] MRCP provides non-invasive high resolution depiction of the biliary tree and has been validated as a highly accurate reference standard for choledocholithiasis, with pooled sensitivity and specificity exceeding 90%.[12] Current European Society of Gastrointestinal Endoscopy (ESGE) guidelines endorse MRCP or EUS as the preferred diagnostic tools for intermediate-probability patients before therapeutic ERCP.[4]

Despite these well-established limitations, CBD stones are frequently not reported on CT examinations that precede or coincide with an eventual MRCP diagnosis of choledocholithiasis. The aim of this study was to determine the frequency of missed CBD stones on abdominal CT in a cohort of patients with MRCP-confirmed choledocholithiasis, and to compare CT-detected and CT-missed cases with respect to stone size and attenuation, CBD diameter, and CT acquisition protocol characteristics.

Methods

This was a single-center retrospective observational study conducted at Kastamonu Training and Research Hospital. Ethical approval was granted by the Kastamonu University Non-Interventional Clinical Research Ethics Committee (approval no: 2026-18, date 19.02.2026). The requirement for individual informed consent was waived given the retrospective nature of the study and the use of anonymised archival data. All data were handled in accordance with applicable data protection legislation.

Electronic medical records and the picture archiving and communication system were searched to identify patients who underwent MRCP between 1 January 2023 and 31 December 2025 for whom choledocholithiasis was confirmed. Eligibility required age ≥18 years, MRCP findings consistent with CBD stones, and at least one abdominal CT performed within seven days before or after MRCP in DICOM format. Patients were excluded if ERCP or surgical stone extraction had been performed between CT and MRCP, if biliary stenting or an artefact precluded CBD assessment, or if image quality was insufficient. The full criteria are summarised in Table 1. Thirty-three patients constituted the final cohort.

All CT examinations were performed for routine clinical indications on a 64-detector, 128-slice scanner (Revolution EVO; GE Healthcare, Chicago, IL, USA). Both contrast-enhanced (portal venous phase) and unenhanced acquisitions were represented. Axial images were acquired at 1.25 mm; multiplanar reconstructions were reviewed when available. MRCP was performed on a 1.5 Tesla system (SIGNA Victor; GE Healthcare). MRCP served as the reference standard.

CT images were independently reevaluated by two blinded radiologists with 1 and 11 years of abdominal imaging experience, respectively. Discordant readings were resolved by consensus. CT positivity was defined as direct visualization of a CBD stone, either as a hyperdense intraluminal focus or as a filling defect (target/crescent sign). A missed CBD stone was defined as a case in which MRCP confirmed choledocholithiasis and retrospective CT rereading identified findings compatible with a stone that had not been reported on the original interpretation.

Statistical Analysis

All analyses were performed using IBM SPSS Statistics (version 26.0). Normality was assessed using the Shapiro-Wilk test. Results are reported as mean ± standard deviation (SD) or median [interquartile range (IQR)] as appropriate. CT sensitivity and 95% confidence intervals (CIs) were calculated using the Wilson method. Continuous variables were compared using the Mann-Whitney U test, and categorical variables were compared using Fisher's exact test. MRCP and CT CBD diameter correlation was assessed by Spearman coefficient; agreement by Bland Altman analysis.[12] The association between stone diameter (≤5 mm vs. >5 mm) and CT non-detection was evaluated by Fisher exact test with odds ratio (OR) and 95% CI. $p<0.05$ was considered statistically significant.

Results

The cohort comprised 33 patients (17 men, 16 women; mean age, 68.1±17.3 years; range, 32-93 years). Baseline characteristics are shown in Table 2. Intrahepatic ductal dilatation was present in 27 patients (81.8%). Contrast-enhanced CT was performed in 20 (60.6%) patients and unenhanced CT in 13 (39.4%) patients. Periduodenal gas was identified on CT in 7 patients (21.2%). Median MRCP stone count was 1 (IQR, 1-2), median stone diameter was 7.0 mm (IQR, 4.0-9.0), and mean MRCP CBD diameter was 11.7±3.7 mm.

CT sensitivity was 48.5% (16/33; 95% CI, 32.5-64.8%), corresponding to a missed finding rate of 51.5% (17/33; 95% CI, 35.2-67.5%). Specificity, positive predictive value, and negative predictive value could not be derived because the cohort consisted exclusively of MRCP-positive patients.

Group comparisons are shown in Table 3. MRCP stone diameter was significantly larger in the CT-detected group [median: 9.0 mm (IQR, 7.8-11.2) vs. 5.0 mm (IQR, 3.0-7.0); U=232; p=0.001]. CT-measured CBD diameter was likewise larger [median, 14.0 mm (IQR, 10.8-16.0) vs. 10.0 mm (IQR, 8.0-12.0); U=202; p=0.017]. Stone count, MRCP CBD diameter, and age did not differ significantly between groups. Intrahepatic ductal dilatation (OR, 2.15; p=0.656), contrast-enhanced CT (OR, 1.17; p=1.000), and periduodenal gas (OR, 3.41; p=0.225) were comparable between groups.

Table 1. Patient inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Age ≥18 years	MRCP to CT interval >7 days
MRCP confirmed choledocholithiasis within the study period	Endoscopic (ERCP) or surgical stone extraction performed between CT and MRCP
At least one abdominal CT examination available within ±7 days of MRCP	Biliary stent, metallic implant, or dense artefact precluding biliary assessment
CT images accessible in DICOM format via PACS	Inadequate image quality, incomplete series, or corrupted DICOM data
	MRCP findings attributable solely to stricture or periampullary neoplasm without concurrent stones

CBD: Common bile duct, CT: Computed tomography, DICOM: Digital Imaging and Communications in Medicine, ERCP: Endoscopic retrograde cholangiopancreatography, MRCP: Magnetic resonance cholangiopancreatography, PACS: Picture archiving and communication system

Table 2. Baseline clinical and imaging characteristics (n=33)	
Variable	**All patients (n=33)**
Age (years), mean ± SD	68.1±17.3
Male sex, n/N (%)	17/33 (51.5%)
MRCP stone count, median (IQR)	1 (1-2); range 1-6
MRCP largest stone diameter (mm), median(IQR)	7.0 (4.0-9.0); range 3-35
MRCP CBD diameter (mm), mean ± SD	11.7±3.7; range 6-20
Intrahepatic ductal dilatation, n (%)	27 (81.8%)
Contrast enhanced CT, n (%)	20 (60.6%)
Unenhanced CT, n (%)	13 (39.4%)
Periduodenal gas on CT, n (%)	7 (21.2%)

Data are mean ± SD or median (IQR) unless stated otherwise
CBD: Common bile duct, IQR: Interquartile range, MRCP: Magnetic resonance cholangiopancreatography, SD: Standard deviation

MRCP and CT CBD diameters were strongly correlated (Spearman r=0.793; p<0.0001). Bland-Altman analysis showed a mean bias of +0.18 mm, SD of differences of 2.34 mm, and 95% LoA of -4.40 to +4.76 mm (Figure 1).

Stratification by MRCP stone diameter showed that CT missed stones ≤5 mm in 10 of 12 patients (83.3%) versus 7 of 21 patients (33.3%) with stones >5 mm (OR, 10.00; p=0.010) (Figure 2).

Among the 16 CT- detected stones, the median attenuation was 78 hounsfield units (HU) (IQR, 56-114 HU; range, 13-505 HU; Shapiro-Wilk p<0.001). Eight stones (50.0%) showed mixed or partially calcified attenuation (20-99 HU), 6 (37.5%) were densely calcified (≥100 HU), and 2 (12.5%) were near isoattenuating (<20 HU) (Table 4, Figure 3-5).

Discussion

The principal finding of this study is that CT failed to identify CBD stones in more than half of the patients with MRCP-confirmed choledocholithiasis, yielding an overall sensitivity of 48.5%. This figure falls within the range reported in the published literature, toward the lower end. Anderson et al.[7] found a sensitivity of 77.2% for multidetector CT in a dedicated prospective series, while Soto et al.[8] reported sensitivities as low as 46% for unenhanced helical CT. A more recent comparative analysis by Al-Dulaimi et al.[11] confirmed that CT performs substantially worse than magnetic resonance imaging/MRCP for CBD stone detection in a real world emergency cohort. The comparatively low sensitivity observed in the present study is consistent with real- world retrospective series and likely attributable to the stone characteristics of the study population, in which small and mixed-attenuation calculi predominated, rather than to systematic technical inadequacy. The heterogeneity across published estimates underscores that CT sensitivity for choledocholithiasis is not a fixed diagnostic property but is heavily dependent on stone characteristics, patient selection, and protocol optimization.

Stone diameter on MRCP emerged as the most robust discriminator between CT-detected and CT-missed cases. The CT-detected group had a significantly larger median stone diameter (9.0 mm vs. 5.0 mm; p=0.001), and when a clinically relevant threshold of 5 mm was applied, the odds of CT non-detection were ten times higher for small stones (OR, 10.00; p=0.010). This finding is consistent with the well established principle that CBD stones smaller than 5 mm are substantially harder to identify on axial CT owing to partial volume averaging and the limited attenuation contrast between small calcifications and surrounding bile at standard reconstruction thicknesses.[7,10] Moreover, small CBD stones carry significant clinical risk microlithiasis is a recognised precipitant of acute biliary pancreatitis and cholangitis even in the absence of substantial ductal dilatation[3,13,14] making their CT invisibility a clinically consequential limitation. These results support a low threshold for MRCP referral in patients with symptoms consistent with biliary obstruction, in whom CT shows no direct evidence of stones but cannot confidently exclude small calculi.

The attenuation profile of CT-detected stones was predominantly mixed or partially calcified, with a median of 78 HU (IQR, 56-114 HU). Only 37.5% of CT visible stones were densely calcified (≥100 HU), while 12.5% exhibited near isoattenuating characteristics. This distribution reflects the known compositional heterogeneity of CBD stones, which range from purely cholesterol (low attenuation, often isoattenuating with bile) to pigment dominant varieties.[7,8] The two near isoattenuating stones identified on CT rereading represent cases in which the diagnosis was recoverable only through systematic scrutiny, careful biliary windowing, and attention to subtle intraluminal filling defects. The 17 stones missed entirely in the original reporting presumably included a disproportionate number of isoattenuating or very small calcifications not visible under standard reporting conditions, which reinforces the conclusion that the diagnostic gap reflects both a physical limitation of CT and a partially addressable reporting behaviour.

Table 3. Comparison of CT detected and CT missed groups			
Variable	**CT detected (n=16)**	**CT missed (n=17)**	**p value**
Continuous variables: Mann-Whitney U test, median (IQR)			
MRCP stone diameter (mm)	9.0 (7.8-11.2)	5.0 (3.0-7.0)	**0.001**
CT CBD diameter (mm)	14.0 (10.8-16.0)	10.0 (8.0-12.0)	**0.017**
MRCP CBD diameter (mm)	11.5 (10.0-15.0)	11.0 (9.5-12.0)	0.338
MRCP stone count	1.0 (1.0-3.0)	1.0 (1.0-2.0)	0.439
Age (years)	63.5 (53.2-78.0)	74.5 (61.2-82.5)	0.377
Categorical variables: Fisher exact test, n (%)			
Intrahepatic ductal dilatation	14 (87.5%)	13 (76.5%)	0.656
Contrast enhanced CT	10 (62.5%)	10 (58.8%)	1.000
Periduodenal gas on CT	5 (31.3%)	2 (11.8%)	0.225

Significant p values are shown in bold. Continuous variables reported as median (IQR)
CBD: Common bile duct, IQR: Interquartile range, MRCP: Magnetic resonance cholangiopancreatography, CT: Computed tomography

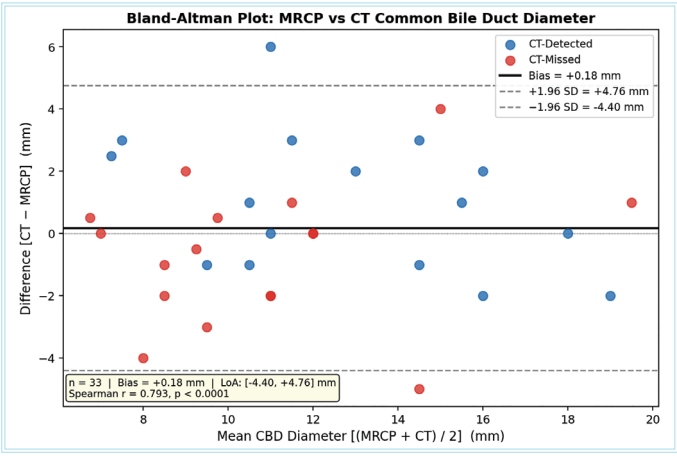


Figure 1. Bland-Altman plot for MRCP and CT CBD diameter agreement (n=33). Solid line=mean bias (+0.18 mm); dashed lines=95% LoA (-4.40 to +4.76 mm). Blue circles, CT detected; red circles, CT missed

CBD: Common bile duct, LoA: Limits of agreement, MRCP: Magnetic resonance cholangiopancreatography, CT: Computed tomography

The CT-measured CBD diameter was significantly larger in the CT-detected group (median, 14.0 mm vs. 10.0 mm; p=0.017). A plausible physiological explanation is that a larger stone is more likely to cause upstream biliary dilatation, increasing duct caliber and improving the conspicuity of an intraluminal filling defect. Conversely, smaller stones may cause only mild or intermittent dilatation. MRCP and CT CBD diameters were strongly correlated (Spearman r=0.793; p<0.0001), and Bland-Altman analysis demonstrated a negligible mean bias of +0.18 mm with 95% LoA of -4.40 to +4.76 mm.[12] Although the wide limits of agreement indicate that individual measurements may deviate by up to approximately 5 mm, the absence of systematic bias confirms that CT reliably tracks CBD diameter at the population level. The broad LoA most likely reflects differences in measurement planes, distension state, and respiratory phase, rather than any fundamental inaccuracy of CT choledochometry.

Neither the use of contrast-enhanced CT (OR, 1.17; p=1.000) nor the presence of periduodenal gas (OR, 3.41; p=0.225) was associated with a significantly higher CT detection rate. These null results must be interpreted with caution, given the limited statistical power of this pilot study. The existing literature suggests that portal venous phase CT can meaningfully improve stone detection by optimising the attenuation contrast between CBD contents and surrounding structures.[10] Systematic use of multiplanar coronal and sagittal reformations has similarly been shown to improve detection of biliary and abdominal pathology compared with axial-only review.[15-19] The failure to detect protocol-related effects in the present series most likely reflects an insufficient sample size rather than a genuine absence of effect, and should not be construed as evidence that contrast enhancement and MPR review are unimportant in biliary CT evaluation.

The rate of missed findings documented here reflects a combination of physical imaging limitations and reporting practices. CBD stones represent one of the more common sources of missed or delayed diagnoses on abdominal CT, partly because the interpreter's attention is directed toward the primary clinical question and the biliary tree receives only secondary scrutiny, a phenomenon analogous to satisfaction of search errors described in other CT contexts.[20] The disparity in experience between the two rereading radiologists (1 year vs. 11 years) reflects the heterogeneity of expertise in real clinical settings.[21] In the present consensus-based design, the potential negative influence of inexperience was mitigated by joint resolution of discordant cases, but in routine single-reader reporting, which characterises most emergency CT workloads, the risk of perceptual oversight remains a real, modifiable component of the diagnostic gap.

The clinical implications of these findings are relevant to several intersecting practice decisions. Current ESGE guidelines recommend MRCP or EUS for patients with intermediate pretest probability of choledocholithiasis before proceeding to ERCP.[4] EUS has demonstrated high sensitivity and specificity for CBD stones in systematic reviews, with particular utility for small stones and borderline ductal dilatation.[13-17] The results of the present study support the view that a CT showing no direct stone evidence should not confidently exclude choledocholithiasis,

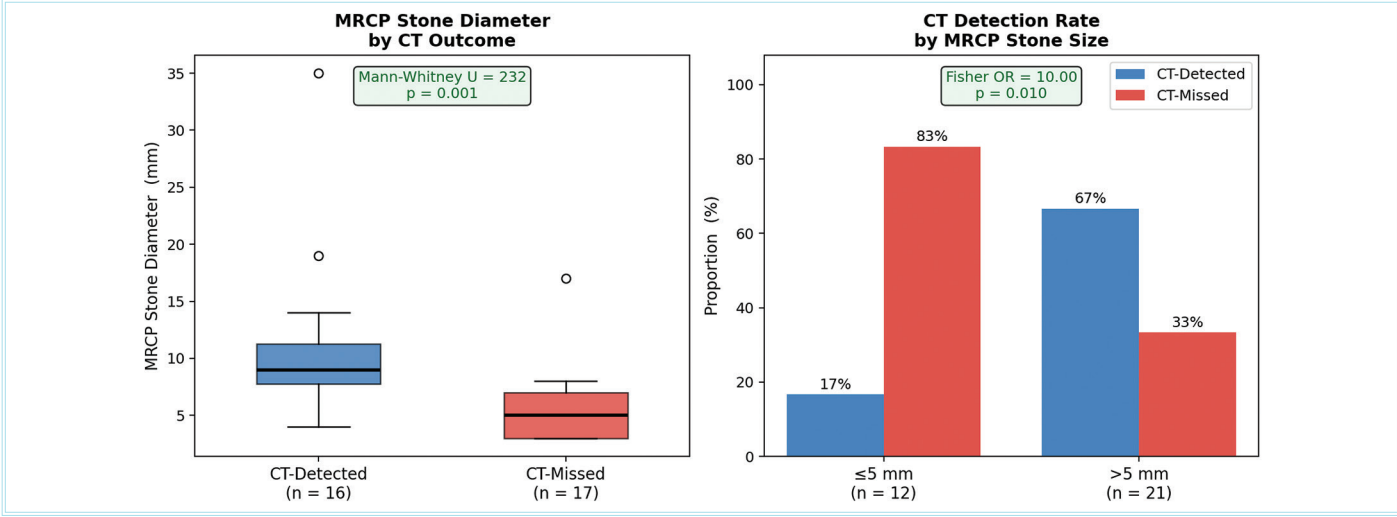


Figure 2. Left: MRCP stone diameter in CT detected (blue, n=16) vs. CT missed (red, n=17) groups (U=232; p=0.001). Right: CT detection rate by stone size (≤5 mm vs. >5 mm; OR=10.00; p=0.010)

MRCP: Magnetic resonance cholangiopancreatography, CT: Computed tomography, OR: Odds ratio

particularly in patients with stones ≤5 mm, mildly dilated ducts around 10 mm, or a syndrome consistent with biliary obstruction. In these cases, proceeding directly to MRCP or EUS where available is likely to reduce diagnostic delay, prevent unnecessary repeat imaging, and facilitate timely biliary drainage.[4,14]

Study Limitations

This study has several limitations. The sample size of 33 patients is small, and the study should be considered a pilot and hypothesis-generating study rather than definitive. The absence of an MRCP negative control group precluded the calculation of specificity, positive predictive value, or negative predictive value. The retrospective single-centre design limits generalisability. Formal interobserver kappa statistics were not calculated as a primary endpoint, as the consensus design precluded separate analysis of individual reader performance; future prospective studies should incorporate inter rater agreement as a designated outcome.[21] The heterogeneous CT protocols reflect real- world practice, but limit protocol- specific subgroup analyses. Finally, laboratory covariates prespecified in the original study design were not available in a sufficiently complete form for inclusion. Despite these limitations, the study provides quantitative real-world data on CT missed-finding rates in MRCP-confirmed choledocholithiasis and identifies stone size as the primary determinant of CT detectability, generating specific hypotheses for adequately powered prospective investigation.

Table 4. CT stone attenuation in the CT detected subgroup (n=16)

Attenuation category	n	%
Densely calcified (≥100 HU)	6	37.5
Mixed/partially calcified (20-99 HU)	8	50.0
Near isoattenuating (<20 HU)	2	12.5
Median attenuation, HU (IQR; range)	**78 (56-114; 13-505)**	

HU: Hounsfield units, IQR: Interquartile range, CT: Computed tomography

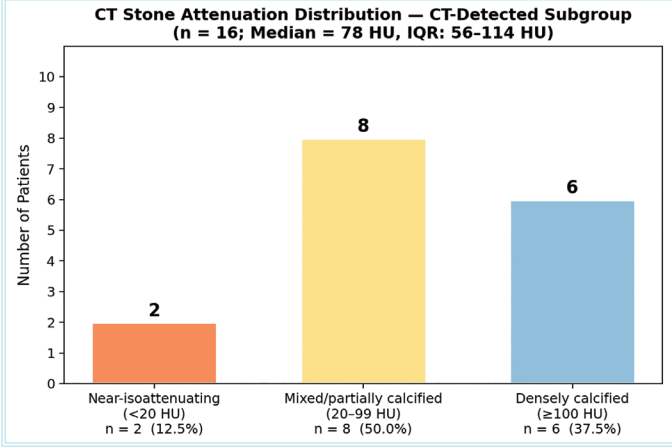


Figure 3. CT stone attenuation distribution in the CT detected subgroup (n=16). Median=78 HU (IQR, 56-114 HU)

HU: Hounsfield units, IQR: Interquartile range, CT: Computed tomography

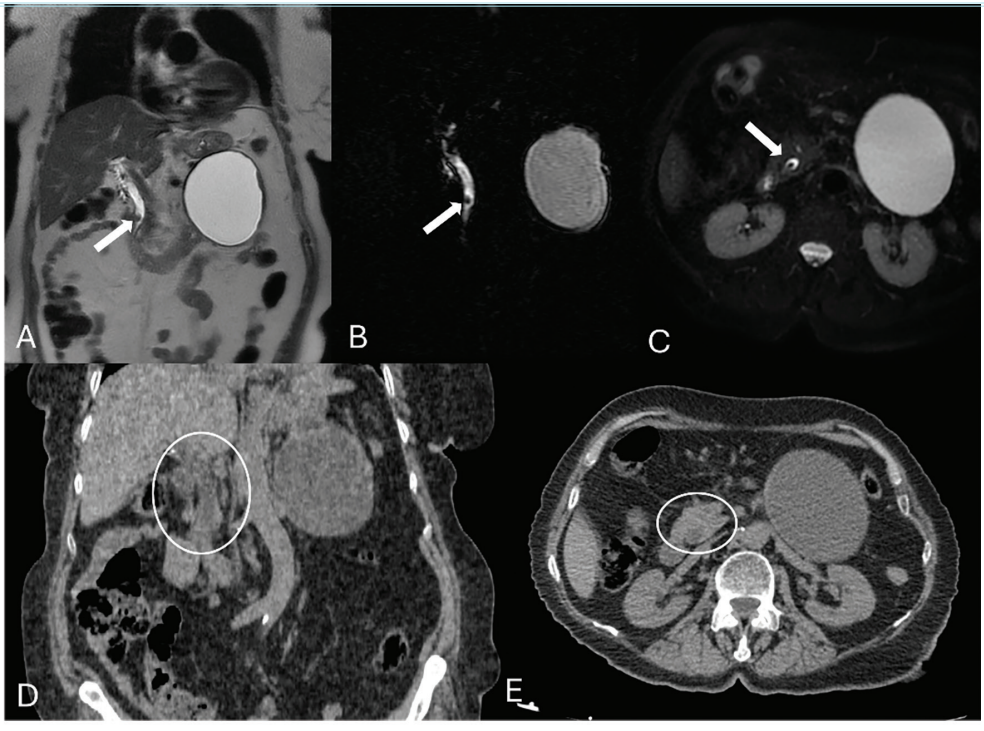


Figure 4. Small distal common bile duct (CBD) stone detected on magnetic resonance cholangiopancreatography but missed on unenhanced computed tomography (CT). Coronal fat suppressed T2-weighted (A), coronal three dimensional magnetic resonance cholangiopancreatography (B), and axial fat suppressed T2-weighted (C) images demonstrate a 4.5 mm hypointense filling defect at the distal CBD (arrows), surrounded by hyperintense bile, consistent with choledocholithiasis. Coronal (D) and axial (E) unenhanced CT images at the corresponding level show a mildly dilated CBD without an identifiable hyperdense focus, filling defect, or target sign; the stone was not reported on the original clinical interpretation or identified on retrospective rereading

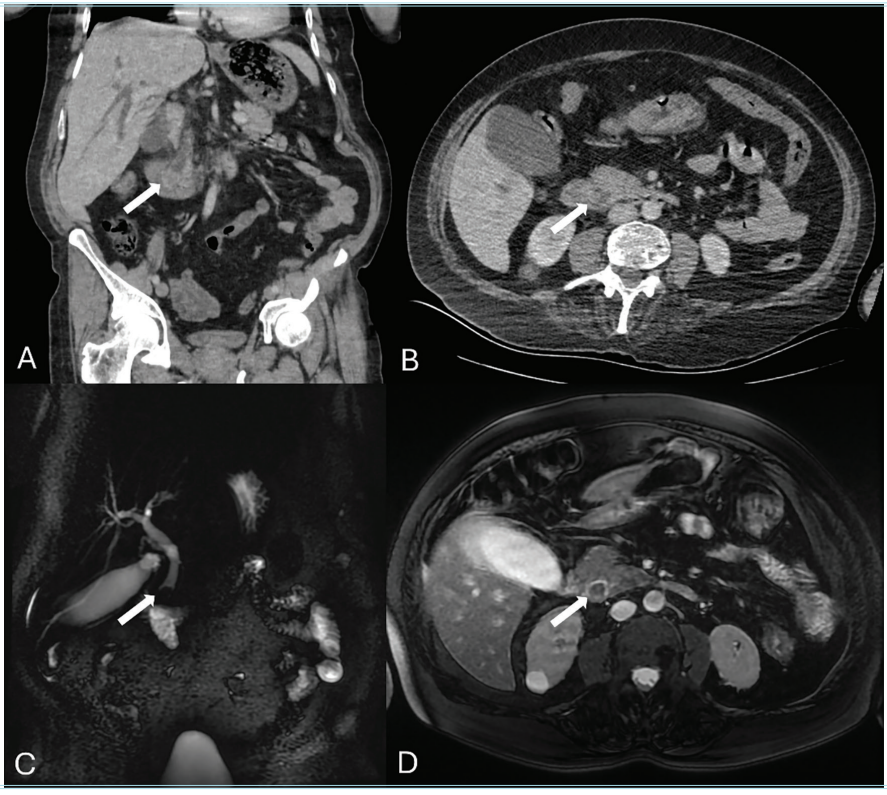


Figure 5. Isoattenuating common bile duct (CBD) stone detected on contrast enhanced CT and confirmed on MRCP. Coronal (A) and axial (B) contrast enhanced CT images demonstrate a 12 mm near isoattenuating intraluminal filling defect within a dilated CBD (arrows), identifiable as a subtle soft tissue density structure surrounded by contrast opacified bile. Coronal radial two dimensional thick slab magnetic resonance cholangiopancreatography (C) and axial fat suppressed T2-weighted (D) images confirm a well defined hypointense filling defect at the distal CBD (arrows), consistent with choledocholithiasis

Conclusion

CT failed to identify CBD stones in more than half of patients with MRCP confirmed choledocholithiasis in this single centre retrospective cohort, yielding a sensitivity of 48.5%. Stone diameter ≤5 mm was the strongest predictor of CT non-detection, with an OR of 10.00 compared with stones larger than 5 mm. MRCP and CT measurements of CBD diameter were strongly correlated, with negligible systematic bias, supporting the reliability of CT choledochometry at the population level. These results reinforce current guideline recommendations in favour of early MRCP referral for patients with a clinical presentation consistent with biliary obstruction, in whom CT does not directly demonstrate a CBD stone. Larger prospective multicentre studies are warranted to confirm these findings, to evaluate the independent contribution of CT protocol factors such as phase selection and slice thickness, and to develop evidence-based reporting standards for biliary CT evaluation.

Ethics

Ethics Committee Approval: Ethical approval was granted by the Kastamonu University Non-Interventional Clinical Research Ethics Committee (approval no: 2026-18, date 19.02.2026).

Informed Consent: Since the study was a retrospective study, informed consent was not required by the ethics committee.

Footnotes

Authorship Contributions

Concept: B.R, F.E.Y., Design: B.R., F.E.Y., Data Collection or Processing: B.R., F.E.Y., Analysis or Interpretation: B.R., F.E.Y., Literature Search: B.R., F.E.Y., Writing: B.R., F.E.Y.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Stinton LM, Shaffer EA. Epidemiology of gallbladder disease: cholelithiasis and cancer. Gut Liver. 2012;6:172-87.
2. ASGE Standards of Practice Committee; Maple JT, Ben-Menachem T, Anderson MA, et al. The role of endoscopy in the evaluation of suspected choledocholithiasis. Gastrointest Endosc. 2010;71:1-9.
3. Kimura Y, Takada T, Kawarada Y, et al. Definitions, pathophysiology, and epidemiology of acute cholangitis and cholecystitis: Tokyo Guidelines. J Hepatobiliary Pancreat Surg. 2007;14:15-26.
4. Manes G, Paspatis G, Aabakken L, et al. Endoscopic management of common bile duct stones: European Society of Gastrointestinal Endoscopy (ESGE) guideline. Endoscopy. 2019;51:472-91.

5. Gurusamy KS, Giljaca V, Takwoingi Y, et al. Ultrasound versus liver function tests for diagnosis of common bile duct stones. Cochrane Database Syst Rev. 2015;2015:CD011548.

6. Gans SL, Pols MA, Stoker J, Boermeester MA; expert steering group. Guideline for the diagnostic pathway in patients with acute abdominal pain. Dig Surg. 2015;32:23-31.

7. Anderson SW, Lucey BC, Varghese JC, Soto JA. Accuracy of MDCT in the diagnosis of choledocholithiasis. AJR Am J Roentgenol. 2006;187:174-80.

8. Soto JA, Alvarez O, Múnera F, Velez SM, Valencia J, Ramírez N. Diagnosing bile duct stones: comparison of unenhanced helical CT, oral contrast-enhanced CT cholangiography, and MR cholangiography. AJR Am J Roentgenol. 2000;175:1127-34

9. Romagnuolo J, Bardou M, Rahme E, Joseph L, Reinhold C, Barkun AN. Magnetic resonance cholangiopancreatography: a meta-analysis of test performance in suspected biliary disease. Ann Intern Med. 2003;139:547-57.

10. Anderson SW, Rho E, Soto JA. Detection of biliary duct narrowing and choledocholithiasis: accuracy of portal venous phase multidetector CT. Radiology. 2008;247:418-27.

11. Al-Dulaimi M, Ibraheem M, Abdulkareem M, Al-Dujaili A, 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.

12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307-10.

13. de Lédinghen V, Lecesne R, Raymond JM, et al. Diagnosis of choledocholithiasis: EUS or magnetic resonance cholangiography? A prospective controlled study. Gastrointest Endosc. 1999;49:26-31.

14. Tenner S, Baillie J, DeWitt J, Vege SS; American College of Gastroenterology. American College of Gastroenterology guideline: management of acute pancreatitis. Am J Gastroenterol. 2013;108:1400-15. Erratum in: Am J Gastroenterol. 2014;109:302.

15. Macari M, Megibow AJ, Balthazar EJ. A pattern approach to the abnormal small bowel: observations at MDCT and CT enterography. AJR Am J Roentgenol. 2007;188:1344-55.

16. Nino-Murcia M, Jeffrey RB Jr, Beaulieu CF, Li KC, Rubin GD. Multidetector CT of the pancreas and bile duct system: value of curved planar reformations. AJR Am J Roentgenol. 2001;176:689-93.

17. Giljaca V, Gurusamy KS, Takwoingi Y, et al. Endoscopic ultrasound versus magnetic resonance cholangiopancreatography for common bile duct stones. Cochrane Database Syst Rev. 2015;2015:CD011549.

18. Kondo S, Isayama H, Akahane M, et al. Detection of common bile duct stones: comparison between endoscopic ultrasonography, magnetic resonance cholangiography, and helical-computed-tomographic cholangiography. Eur J Radiol. 2005;54:271-5.

19. Catalano OA, Sahani DV, Forcione DG, et al. Biliary infections: spectrum of imaging findings and management. Radiographics. 2009;29:2059-80.

20. Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. Ulster Med J. 2012;81:3-9.

21. Mahgerefteh S, Kruskal JB, Yam CS, Blachar A, Sosna J. Peer review in diagnostic radiology: current state and a vision for the future. Radiographics. 2009;29:1221-31.

RESEARCH ARTICLE

DOI: 10.4274/AdvRadiolImaging.galenos.2026.58077
Adv Radiol Imaging 2026;3(1):15-24

Correlation of CT Findings with Pathology and Interobserver Agreement in Patients Undergoing Appendectomy for Suspected Acute Appendicitis: A Single-center Retrospective Study

Erdem Özkan[1], Abdullah Can[1], Nurtaç Sarıkaş[2], Yasin Alper Yıldız[3]

[1]*Kastamonu Training and Research Hospital, Clinic of Radiology, Kastamonu, Türkiye*
[2]*Kastamonu University Faculty of Medicine, Department of Pathology, Kastamonu, Türkiye*
[3]*Kastamonu University Faculty of Medicine, Department of General Surgery, Kastamonu, Türkiye*

Abstract

Objectives: Computed tomography (CT) has become the cornerstone imaging modality for the evaluation of suspected acute appendicitis, yet real-world diagnostic performance data from single-centre experiences incorporating both CT-pathology correlation and interobserver agreement remain limited. This study aimed to evaluate the diagnostic performance of CT, using histopathology as the reference standard, to characterise individual CT imaging features, and to assess interobserver reproducibility in a consecutive adult cohort.

Methods: A retrospective analysis was conducted at our institution involving 195 adult patients who underwent appendectomy after preoperative contrast-enhanced abdominal CT for clinically suspected acute appendicitis between January 2023 and February 2026. Histopathological examination served as the reference standard. Diagnostic performance metrics (sensitivity, specificity, positive and negative predictive values, overall accuracy) were calculated from the 2×2 contingency table, with 95% confidence intervals (CIs) estimated using the Wilson score method. Odds ratios (ORs) with 95% CIs were computed for each CT feature using Fisher's exact test. Receiver operating characteristic (ROC) analysis with the Youden-optimal cut-point was performed for the appendiceal diameter. Interobserver agreement for CT features was assessed in a 40-patient subsample evaluated by two radiologists with differing levels of experience, using Cohen's kappa and the intraclass correlation coefficient (ICC).

Results: Acute appendicitis was histopathologically confirmed in 155 patients (79.5%). CT demonstrated a sensitivity of 98.7% (95% CI: 95.4-99.6%), a specificity of 87.5% (73.9-94.5%), a positive predictive value of 96.8% (92.8-98.6%), a negative predictive value of 94.6% (82.3-98.5%), and an overall accuracy of 96.4% (92.8-98.3%). Forty appendectomy specimens (20.5%) showed no histopathological evidence of acute appendicitis and were classified as the pathology-negative group ($Path^-$). Periappendiceal fat stranding demonstrated the strongest independent association with confirmed appendicitis (OR: 173.83; 95% CI: 50.17-602.28; $p<0.001$), followed by contrast enhancement (OR: 120.00; 95% CI: 36.85-390.82; $p<0.001$) and wall thickening (OR: 49.23; 95% CI: 17.92-135.22; $p<0.001$). The presence of an appendicolith was not significantly associated with confirmed appendicitis ($p=0.151$). ROC analysis of the appendiceal diameter yielded an area under the ROC curve of 0.839, with an optimal cut point of 8.3 mm (sensitivity 86.5%, specificity 75.0%). All five false-positive cases were attributable to lymphoid hyperplasia. Both false-negative cases involved atypical presentations-one with early-stage disease and the other gangrenous appendicitis with an attenuated inflammatory response. Interobserver agreement for the overall CT diagnosis was perfect (Cohen's kappa coefficient=1.000), and appendiceal diameter measurement demonstrated excellent reproducibility (ICC=0.994).

Conclusion: CT achieves high diagnostic accuracy for acute appendicitis in real-world clinical practice. Periappendiceal fat stranding is the single most predictive CT finding, while the presence of an appendicolith is not independently statistically significant. Gangrenous appendicitis and early-stage disease are the principal causes of CT false-negative results. Interobserver agreement for the overall diagnosis is excellent, although experience-dependent variability persists for individual secondary signs.

Keywords: Acute appendicitis, computed tomography, diagnostic performance, interobserver agreement

Cite this article as: Özkan E, Can A, Sarıkaş N, Yıldız YA. Correlation of CT findings with pathology and interobserver agreement in patients undergoing appendectomy for suspected acute appendicitis: a single-center retrospective study. Adv Radiol Imaging. 2026;3(1):15-24

Address for Correspondence: Erdem Özkan MD, Kastamonu Training and Research Hospital, Clinic of Radiology, Kastamonu, Türkiye
E-mail: erdemozkan5454@gmail.com **ORCID ID:** orcid.org/0000-0001-8120-7051
Received: 08.03.2026 **Accepted:** 22.04.2026 **Published:** 30.04.2026

Introduction

Acute appendicitis is the most prevalent cause of emergency abdominal surgery worldwide, with an incidence of approximately 100 cases per 100,000 persons per year and a lifetime risk estimated at 7-9% in developed countries.[1,2] Despite centuries of clinical experience, the diagnosis of acute appendicitis remains a significant challenge in daily practice. Clinical evaluation, supported by laboratory findings and scoring systems such as the Alvarado score, does not reliably achieve sufficient accuracy to guide surgical decision-making, particularly in patients with atypical presentations, fertile women, elderly patients, and immunosuppressed individuals.[3,4,5]

Contrast-enhanced computed tomography (CT) of the abdomen and pelvis has transformed the preoperative evaluation of suspected acute appendicitis, and is currently recommended as the initial imaging modality of choice for non-pregnant adults by both the 2024 Infectious Diseases Society of America (IDSA) guidelines and the 2020 World Society of Emergency Surgery Jerusalem guidelines.[6,7] A landmark Cochrane systematic review encompassing 64 studies and 10,280 participants reported a pooled sensitivity of approximately 95% and specificity of approximately 94% for CT in this clinical context,[1] results broadly consistent with large single-centre experiences demonstrating sensitivities above 98% in high-volume multidetector CT practice.[8]

The fundamental CT findings associated with acute appendicitis have been extensively characterised and include appendiceal distension, wall thickening, contrast enhancement, periappendiceal fat stranding, free fluid, appendicolith, and CT-detected perforation.[9,10,11] Findings carry distinct diagnostic weight, and their clinical interpretation is further complicated by the well-recognised entity of lymphoid hyperplasia-a benign reactive condition capable of producing an appendiceal CT appearance indistinguishable from early appendicitis, particularly in younger patients.[12,13] The presence of an appendicolith has attracted renewed interest in the era of antibiotic-first management, as converging evidence suggests that appendicolith-associated appendicitis carries a significantly higher risk of complicated disease and failure of non-operative treatment.[14,15,16]

Notwithstanding the overall high accuracy of CT, residual diagnostic uncertainty persists. False-negative CT examinations, although uncommon, primarily occur in two settings: early-stage appendicitis with an appendiceal diameter below the commonly applied threshold and gangrenous appendicitis in which transmural necrosis paradoxically attenuates the inflammatory imaging response.[17,18] The negative appendectomy rate—reflecting cases in which appendectomy reveals a histologically normal appendix or an appendix without appendicitis—remains an important quality metric, with contemporary CT-guided series reporting rates of 5-20%.[4,19] CT-detected perforation is a critical but technically challenging diagnosis, with substantial variability in both its definition and the degree of experience-dependent interobserver agreement.[20,21]

Interobserver agreement for CT interpretation in suspected appendicitis has been a source of ongoing concern. Several studies have demonstrated experience-dependent variability in the identification of individual CT signs, whereas agreement for the overall categorical diagnosis of appendicitis tends to be higher.[20] Understanding the magnitude and sources of this variability has important implications for radiological training programmes and the safe implementation of CT-based diagnostic pathways, particularly in centres where specialist radiological expertise is not available around the clock.

Emerging technologies, including artificial intelligence-assisted CT interpretation, hold promise for further improving diagnostic performance and reducing observer dependency,[22] yet robust local validation data remain essential before such approaches can supplant or substantively augment established radiological workflows.

Against this background, the present study was designed to evaluate the diagnostic performance of preoperative contrast-enhanced CT, with histopathology as the reference standard, in a consecutive adult cohort undergoing appendectomy at a single training and research hospital. Secondary objectives were to determine the diagnostic value of individual CT features, to characterise the clinical and imaging profiles of discordant cases (false-positive and false-negative examinations), and to assess interobserver agreement for CT features across radiologists with differing levels of clinical experience.

Methods

Study Design and Ethical Approval

This single-centre retrospective cross-sectional study was conducted at our hospital in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the Kastamonu University Non-Interventional Clinical Research Ethics Committee (approval number: 2026-49, date: 19.03.2026). Given the retrospective nature of the study, the requirement for informed consent was waived. Patient data were anonymised prior to analysis, and access was restricted to the research team in compliance with applicable data protection regulations. The study was conducted without external funding; all research expenses were covered by the investigators.

Study Population

Medical records and the Picture Archiving and Communication System were retrospectively reviewed from January 1, 2023, to February 1, 2026. Adult patients (≥18 years) who underwent appendectomy following abdominal CT for clinically suspected acute appendicitis were screened for eligibility. The inclusion criteria were: (1) clinical suspicion of acute appendicitis resulting in surgical intervention; (2) availability of a preoperative contrast-enhanced abdominal CT examination; (3) availability of a postoperative histopathological report; and (4) age ≥18 years. Exclusion criteria were: unavailable or non-retrievable CT images; non-visualisation of the vermiform appendix on CT; inadequate image quality due to motion or other artefacts; and incomplete histopathological data. Ultimately, 195 patients met the study criteria and were included in the final analysis.

CT Examination and Image Evaluation

All CT examinations were performed on a 128-Slice Revolution Maxima CT scanner (GE HealthCare) with intravenous contrast enhancement. CT images were independently evaluated by two radiologists with differing levels of experience (reader 1: 1 year; reader 2: 8 years), and were both blinded to patients' clinical information and histopathological results. The following imaging features were recorded using a standardised data collection form: (1) appendiceal diameter (outer-to-outer measurement, mm); (2) appendiceal wall thickening (present/absent); (3) appendiceal wall contrast enhancement (present/absent); (4) periappendiceal fat stranding (present/absent); (5) free fluid in the periappendiceal

region (present/absent); (6) appendicolith (present/absent); and (7) CT-detected perforation (present/absent). The overall CT diagnosis of acute appendicitis (positive or negative) was recorded for each case. In cases of disagreement between the two readers, a consensus decision, reached through joint re-evaluation, was used as the final CT diagnosis for the primary diagnostic performance analysis. Interobserver agreement for each imaging feature was assessed separately using data from the initial independent readings.

Reference Standard

Histopathological examination of the resected appendix specimen served as the reference standard. Acute appendicitis was confirmed when the pathology report demonstrated transmural or mucosal neutrophilic infiltration consistent with acute inflammation. Confirmed cases were further classified as complicated—defined as the presence of perforation and/or gangrene—or uncomplicated. For the purposes of this study, the pathology-negative group ($Path^-$) was defined as appendectomy specimens showing no histopathological evidence of acute appendicitis. This classification was used for CT-histopathology correlation and should not be interpreted as an institutional negative appendectomy rate.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). A two-tailed p value of less than 0.05 was considered statistically significant. The normality of continuous variables was assessed using the Shapiro-Wilk test. Since both age and appendiceal diameter deviated significantly from normality, they are reported as medians with interquartile ranges (IQR). Categorical variables are expressed as absolute numbers and percentages. Between-group comparisons of continuous variables were performed using the Mann-Whitney U test. Associations between categorical CT findings and histopathological outcome were evaluated using Fisher's exact test. The chi-square test was used for sex-stratified comparisons of pathological confirmation rates. Diagnostic performance metrics—sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy—were calculated from the 2×2 contingency table, with 95% confidence intervals (CI) derived using the Wilson score method. Receiver operating characteristic (ROC) curve analysis was performed for the appendiceal diameter; the area under the curve (AUC) was reported, and the optimal diameter threshold was determined by the Youden index (J=sensitivity + specificity -1). For each individual CT feature, odds ratios (ORs) with 95% CIs were calculated from the contingency table using the Haldane-Anscombe correction where applicable. Interobserver agreement for categorical CT features was quantified using Cohen's kappa coefficient (κ), and interobserver agreement for the continuous appendiceal diameter measurement was quantified using the intraclass correlation coefficient (ICC, two-way mixed model, absolute agreement). Kappa values were interpreted according to the scale proposed by Landis and Koch: <0.00, poor; 0.00-0.20, slight; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; and 0.81-1.00, almost perfect.

Results

Patient Demographics and General Characteristics

A total of 195 patients met the inclusion criteria and were enrolled in the study (Table 1). The cohort comprised 99 female patients (50.8%) and 96 male patients (49.2%). The median age was 33 years (IQR: 25.5-44.5; range: 18-73). Both age and appendiceal diameter deviated significantly from normality (Shapiro-Wilk test; $p<0.001$ for both); consequently, non-parametric methods were used for all subsequent analyses. The overall median appendiceal diameter was 10.3 mm (IQR: 8.1-13.2).

Table 1. Demographic characteristics and distribution of histopathological findings (N=195)

Variable	Female (n=99)	Male (n=96)	Total (n=195)	p value
Demographics				
Age - median (IQR), years	33 (26-44)	33 (25-45)	33 (25.5-44.5)	NS
Age range, years	18-73	18-71	18-73	-
Appendiceal diameter - median (IQR), mm	10.1 (7.8-13.0)	10.5 (8.6-13.5)	10.3 (8.1-13.2)	NS
Normality (Shapiro-Wilk)	-	-	Non-normal (p<0.001)	-
Histopathological outcomes				
Confirmed appendicitis ($Path^+$)	71 (71.7%)	84 (87.5%)	155 (79.5%)	0.011
Complicated (perforation and/or gangrene)	-	-	18 (11.6%)[a]	-
Isolated perforation	-	-	9 (5.8%)[a]	-
Gangrene	-	-	9 (5.8%)[a]	-
Abscess/phlegmon	-	-	53 (34.2%)[a]	-
Uncomplicated	-	-	137 (88.4%)[a]	-
Pathology-negative appendectomy specimens ($Path^-$)	28 (28.3%)	12 (12.5%)	40 (20.5%)	-
Normal appendix histology	-	-	16 (40%)[b]	-
Lymphoid hyperplasia	-	-	14 (35.0%)[b]	-
Fibrous obliteration	-	-	9 (22.5%)[b]	-
Congestion	-	-	1 (2.5%)[b]	-

[a]: Percentage of pathology-positive cases (n=155). [b]: Percentage of $Path^-$ cases (n=40). Sex-stratified comparison of pathological confirmation rates performed using the chi-square test. Data are expressed as n (%) unless stated otherwise
$Path^+$: Histopathologically confirmed appendicitis, $Path^-$: Negative appendectomy, IQR: Interquartile range, NS: Not significant (Mann-Whitney U test)

Histopathological Findings

Histopathological examination confirmed acute appendicitis in 155 patients (79.5%), whereas 40 patients (20.5%) had appendectomy specimens without histopathological evidence of acute appendicitis and were therefore classified as the pathology-negative group ($Path^-$) (Table 1). Among the 40 pathology-negative cases, 18 (11.6%) were classified as complicated appendicitis: 9 (5.8%) had isolated perforation and 9 (5.8%) had gangrene. Abscess or phlegmon formation was identified in 53 pathology-positive cases (34.2%). Among the 40 negative appendectomy cases, histopathological findings included: a normal appendix in 16 cases (40%), lymphoid hyperplasia in 14 cases (35.0%), fibrous obliteration in 9 cases (22.5%), and congestion in 1 case (2.5%). The rate of histopathologically confirmed appendicitis was significantly higher among male patients (84/96, 87.5%) than among female patients (71/99, 71.7%) (χ^2 test, p=0.011).

Overall CT Diagnostic Performance

The overall diagnostic performance of CT for the detection of acute appendicitis is summarised in Table 2. CT yielded a sensitivity of 98.7% (95% CI: 95.4-99.6%), a specificity of 87.5% (95% CI: 73.9-94.5%), a PPV of 96.8% (95% CI: 92.8-98.6%), an NPV of 94.6% (95% CI: 82.3-98.5%), and an overall accuracy of 96.4% (95% CI: 92.8-98.3%).

Of 158 CT-positive examinations, 153 (96.8%) were confirmed as true-positive cases. Five CT-positive cases (3.2%) were false positives, representing patients with imaging findings suggestive of appendicitis but without histopathological confirmation of acute appendicitis. Of the 37 CT-negative examinations, 35 (94.6%) were true-negatives. Two CT-negative cases (5.4%) were false-negatives (FNs), representing missed cases of appendicitis.

Diagnostic Value of Individual CT Findings

The diagnostic performance of individual CT features is detailed in Table 3. All primary inflammatory CT signs—periappendiceal fat stranding, contrast enhancement, and wall thickening—were significantly associated with histopathologically confirmed appendicitis (Fisher's exact test, p<0.001 for all three).

Periappendiceal fat stranding showed the strongest independent association with pathological appendicitis (OR: 173.83, 95% CI: 50.17-602.28; p<0.001), and was present in 149 of 155 pathology-positive cases (96.1%) but in only 5 of 40 pathology-negative cases (12.5%), yielding a sensitivity of 96.1%, a specificity of 87.5%, and a PPV of 96.8%. Contrast enhancement demonstrated the highest sensitivity among all individual features (96.8%), with an OR of 120.00 (95% CI: 36.85-390.82; p<0.001). Wall thickening was present in 89.7% of pathology-positive cases, with an OR of 49.23 (95% CI: 17.92-135.22; p<0.001).

Table 2. Overall CT diagnostic performance for acute appendicitis (N=195)

	$Path^+$ (n=155)	$Path^-$ (n=40)	Total
CT positive	TP=153	FP=5	158
CT negative	FN=2	TN=35	37
Total	155	40	195
Diagnostic performance metrics			
	Value	**95% CI (lower)**	**95% CI (upper)**
Sensitivity	98.7%	95.4%	99.6%
Specificity	87.5%	73.9%	94.5%
Positive predictive value	96.8%	92.8%	98.6%
Negative predictive value	94.6%	82.3%	98.5%
Overall accuracy	96.4%	92.8%	98.3%

Sensitivity=TP/(TP+FN), Specificity=TN/(TN+FP), PPV=TP/(TP+FP), NPV=TN/(TN+FN), Accuracy=(TP+TN)/N

CT: Computed tomography, $Path^+$: Histopathologically confirmed appendicitis, $Path^-$: Absence of histopathological evidence of acute appendicitis, TP : True-positive, FP: False-positive, FN: False-negative, TN: True-negative, CI: Confidence interval (Wilson score method)

Table 3. Diagnostic performance of individual CT findings for the detection of acute appendicitis

CT feature	$Path^+$ n/N (%)	$Path^-$ n/N (%)	Sens.	Spec.	PPV	NPV	OR (95% CI) p value
Wall thickening	139/155 (89.7%)	6/40 (15.0%)	89.7%	85.0%	95.9%	68.0%	49.23 (17.92-135.22) p<0.001
Contrast enhancement	150/155 (96.8%)	8/40 (20.0%)	96.8%	80.0%	94.9%	86.5%	120.00 (36.85-390.82) p<0.001
Periappendiceal fat stranding	149/155 (96.1%)	5/40 (12.5%)	96.1%	87.5%	96.8%	85.4%	173.83 (50.17-602.28) p<0.001
Free fluid	74/155 (47.7%)	2/40 (5.0%)	47.7%	95.0%	97.4%	31.9%	17.36 (4.05-74.48) p<0.001
Appendicolith†	41/155 (26.5%)	6/40 (15.0%)	26.5%	85.0%	87.2%	23.0%	2.04 (0.80-5.21) p=0.151
CT-detected perforation†	14/155 (9.0%)	0/40 (0.0%)	9.0%	100.0%	100.0%	22.1%	∞ p=0.078
Appendiceal diameter (ROC; AUC=0.839)*	11.2 mm (9.1-13.6)	6.8 mm (5.5-8.3)	86.5%*	75.0%*	-	-	AUC 0.839 p<0.001

*: Sensitivity and specificity for appendiceal diameter correspond to the Youden-optimal cut-point of 8.3 mm. †: Features that did not reach statistical significance (p≥0.05)- CT-detected perforation: 14/155 cases in $Path^+$ group, 0/40 in $Path^-$ group. Values for categorical features represent frequency (n/N, %). Diameter values represent median (IQR). OR: Odds ratio calculated using Fisher's exact test (two-tailed); 95% CI for OR computed with Haldane-Anscombe correction where applicable

Sens.: Sensitivity, Spec.: Specificity, PPV: Positive predictive value, NPV: Negative predictive value, AUC: Area under the receiver operating characteristic curve, CI: Confidence interval, CT: Computed tomography

Free fluid was identified in 74 pathology-positive cases (47.7%) and in only 2 pathology-negative cases (5.0%), reflecting a high specificity of 95.0% and a PPV of 97.4% (OR: 17.36, 95% CI: 4.05-74.48; p<0.001).

CT-detected perforation, when present, carried a PPV of 100% for pathological appendicitis; all 14 CT-reported perforation cases (9.0% of pathology-positive group) were pathologically confirmed. However, the overall association with histopathologically complicated appendicitis did not reach statistical significance (p=0.078), reflecting the limited sensitivity of CT for detecting perforation compared with the histopathological complication rate.

In contrast, an appendicolith was present in 26.5% of pathology-positive and 15.0% of pathology-negative cases, and did not demonstrate a statistically significant association with confirmed appendicitis (OR: 2.04, 95% CI: 0.80-5.21; p=0.151).

Appendiceal Diameter Analysis

The median appendiceal diameter was significantly larger in pathology-positive cases compared to pathology-negative cases [11.2 mm (IQR: 9.1-13.6) vs. 6.8 mm (IQR: 5.5-8.3); Mann-Whitney U test, p<0.001] (Table 3). ROC curve analysis demonstrated an AUC of 0.839, indicating good discriminatory performance. The Youden-optimal cut-point was 8.3 mm, corresponding to a sensitivity of 86.5% and a specificity of 75.0%.

Analysis of Discordant Cases

The characteristics of the five false-positive (FP) and two FN cases are presented in detail in Table 4a, 4b.

Among the five FP cases, four were female and one was male (median age: 33 years; range: 19-49). All five cases demonstrated periappendiceal fat stranding on CT. Histopathological findings in FP cases included lymphoid hyperplasia in four patients (80.0%) and combined lymphoid hyperplasia with fibrous obliteration in one patient (20.0%). In two cases, the appendiceal diameter exceeded 13 mm, and all secondary CT signs were positive, suggesting that the CT appearance was highly suggestive of appendicitis despite a benign histopathological substrate-likely reflecting reactive periappendiceal inflammatory changes secondary to lymphoid hyperplasia.

Both FN cases were young patients (aged 20 and 25 years). The first case was a 20-year-old male with an appendiceal diameter of 5.4 mm and no secondary CT signs (wall thickening, contrast enhancement, fat stranding, free fluid, or appendicolith), consistent with early-stage appendicitis that was below the CT detection threshold (Figure 1). The second was a 25-year-old female whose diameter was 7.9 mm and who had gangrenous changes on pathology, in the absence of convincing CT findings-a pattern consistent with atypical or gangrenous appendicitis, in which transmural necrosis may paradoxically reduce or attenuate the inflammatory imaging response (Figure 2). Both FN patients were ultimately taken to the operating theatre based on persistent clinical suspicion, and appendicitis was confirmed histopathologically in both cases.

Interobserver Agreement

Interobserver agreement was assessed in a subset of 40 patients who were independently evaluated by two radiologists with differing levels of experience (Reader 1: 1 year; Reader 2: 8 years). Results

Table 4a. False-positive cases-CT positive/pathology negative (n=5)

Case	Sex	Age (year)	Diameter (mm)	Wall thickening	Contrast enhancement	Fat stranding	Free fluid	Appendicolith	Histopathology/remark
1	F	19	8.2	-	-	+	-	-	Lymphoid hyperplasia; fat stranding only; reactive change
2	M	19	10.0	-	+	+	-	+	Lymphoid hyperplasia; appendicolith + reactive enhancement
3	F	32	8.0	+	+	+	-	-	Lymphoid hyperplasia + fibrous obliteration; dual pathology; convincing CT pattern
4	F	49	13.1	+	+	+	+	+	Lymphoid hyperplasia; all secondary signs positive; reactive substrate
5	F	33	13.3	+	+	+	+	+	Lymphoid hyperplasia; dilated diameter + full secondary sign complex

All five false-positive cases were managed with appendectomy; lymphoid hyperplasia was the most frequent underlying histopathological finding (4/5, 80%)
+: Present, -: Absent, F: Female, M: Male

Table 4b. False-negative cases-CT negative/pathology positive (n=2)

Case	Sex	Age (year)	Diameter (mm)	Wall thickening	Contrast enhancement	Fat stranding	Free fluid	Appendicolith	Histopathology/reason for miss
1	M	20	5.3	-	-	-	-	-	Acute appendicitis (mild); below-threshold diameter; no secondary signs → early-stage disease
2	F	25	7.1	-	-	-	-	-	Acute appendicitis + gangrene; atypical presentation; transmural necrosis attenuated inflammatory CT response

Both false-negative patients underwent appendectomy based on persistent clinical suspicion; histopathological diagnosis was confirmed postoperatively in both cases
CT: Computed tomography

are summarised in Table 5. Agreement on the overall CT diagnosis of appendicitis was perfect (κ=1.000; 100% raw concordance), indicating that both readers arrived at identical final diagnoses in all 40 cases. Among individual CT features, periappendiceal fat stranding and appendicolith each demonstrated almost perfect agreement (κ=0.805, 95% CI: 0.541-1.068). Contrast enhancement and wall thickening showed substantial agreement (κ=0.754, 95% CI: 0.486-1.022; κ=0.630, 95% CI: 0.326-0.933, respectively). Agreement for free fluid was moderate (κ=0.504, 95% CI: 0.237-0.770); the majority of disagreements (7 of 10) were attributable to reader 1 scoring free fluid as present when reader 2 did not. CT-detected perforation yielded a near-zero kappa (κ=-0.034); however, raw agreement was 92.5% (37/40), and this discrepancy reflects the well-recognised prevalence-dependency of Cohen's kappa (the "kappa paradox"): the very low prevalence of CT-detected perforation in the subsample (≤10%) renders the statistic artificially low despite high observed concordance. Appendiceal diameter measurements demonstrated excellent reproducibility, with an ICC of 0.994 (95% CI: 0.988-0.997), a mean inter-reader difference of -0.01 mm, and Bland-Altman limits of agreement of -0.79 to +0.76 mm.

Discussion

The present study evaluated the real-world diagnostic performance of preoperative contrast-enhanced CT, with histopathology as the gold standard, in a consecutive cohort of adults undergoing appendectomy for clinically suspected acute appendicitis. Our findings confirm that

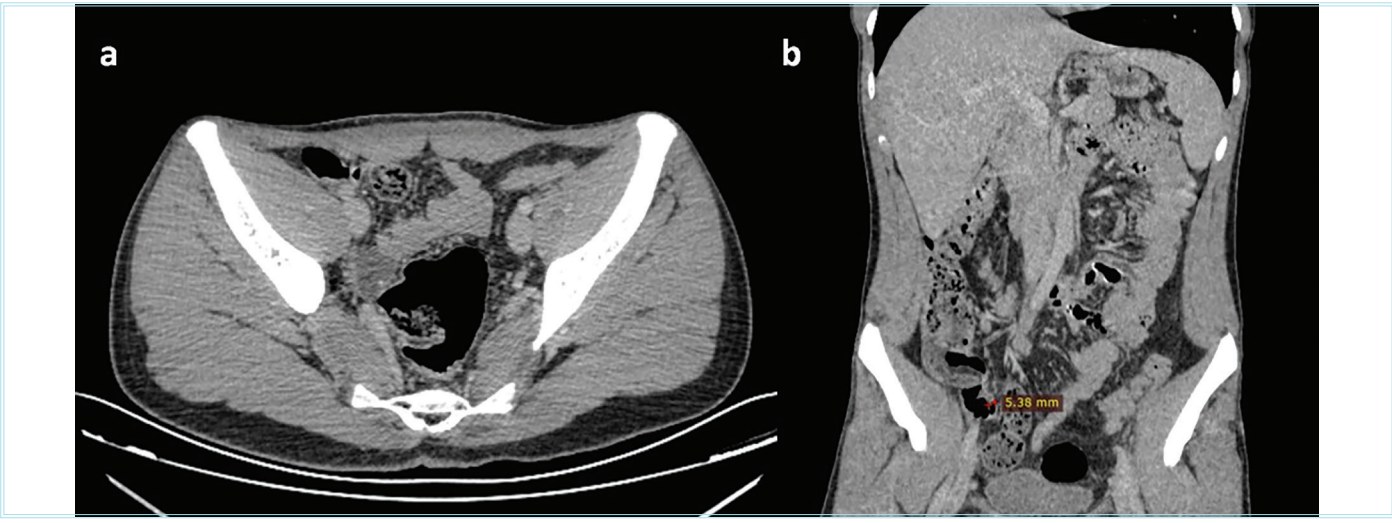


Figure 1. False-negative CT in early-stage appendicitis in a 20-year-old male patient

(a) Axial and (b) coronal CT images demonstrate a non-dilated appendix in the right lower quadrant measuring 5.4 mm, without wall thickening or secondary inflammatory findings such as periappendiceal fat stranding, free fluid, or appendicolith. Although the radiological findings were not typical of appendicitis, histopathological examination revealed early inflammatory changes

CT: Computed tomography

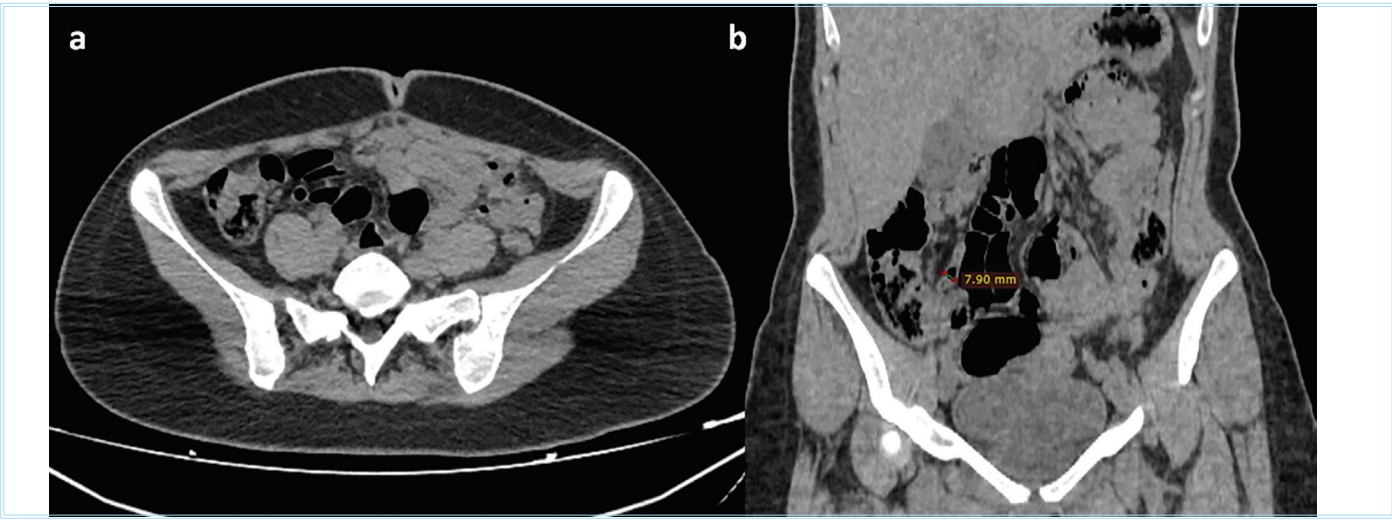


Figure 2. False-negative CT in gangrenous appendicitis in a 25-year-old female patient

(a) Axial and (b) coronal CT images show an appendix measuring 7.9 mm without convincing secondary CT findings of appendicitis. Histopathological analysis revealed gangrenous appendicitis, despite the absence of radiological features typically suggestive of a gangrenous process

CT: Computed tomography

Table 5. Interobserver agreement for CT imaging features (n=40)

CT feature	R1 n (%)	R2 n (%)	Agreement n/40	Agreement (%)	Disagr. $R1^+/R2^-$	Disagr. $R1^-/R2^+$	κ (95% CI)	Strength
Wall thickening	33 (82.5%)	29 (72.5%)	35/40	87.5%	4	1	0.630 (0.326-0.933)	Substantial
Contrast enhancement	35 (87.5%)	33 (82.5%)	37/40	92.5%	2	1	0.754 (0.486-1.022)	Substantial
Periappendiceal fat stranding	34 (85.0%)	32 (80.0%)	38/40	95.0%	2	0	0.805 (0.541-1.068)	Almost perfect
Free fluid	17 (42.5%)	13 (32.5%)	30/40	75.0%	7	3	0.504 (0.237-0.770)	Moderate
Appendicolith	8 (20.0%)	10 (25.0%)	38/40	95.0%	0	2	0.805 (0.541-1.068)	Almost perfect
CT-detected perforation†	3 (7.5%)	4 (10.0%)	37/40	92.5%	1	2	-0.034 (-1.160-1.091)	Poor† (kappa paradox)
Overall CT diagnosis	34 (85.0%)	34 (85.0%)	40/40	100.0%	0	0	1.000 (1.000-1.000)	Perfect
Continuous Variable								
Variable	**R1 Median (IQR), mm**		**R2 Median (IQR), mm**		**Mean diff. (R1-R2)**		**LoA (±1.96 SD)**	**ICC (95% CI)**
Appendiceal diameter	10.1 (8.3-13.0)		10.1 (8.1-12.6)		-0.01 mm		-0.79 to +0.76 mm	0.994 (0.988-0.997)

95% CI computed using the asymptotic standard error method for κ and F-distribution-based method for ICC. Strength of agreement for κ: <0.00 poor; 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect (Landis & Koch, 1977). †: The near-zero kappa for CT-detected perforation reflects the kappa paradox (prevalence-dependency): raw agreement was 92.5% (37/40), but the very low prevalence of perforation (≤10%) renders kappa artificially low despite high observed concordance
κ: Cohen's kappa coefficient, ICC: Intraclass correlation coefficient (two-way mixed model, absolute agreement), R1: Reader 1 (1 year of experience), R2: Reader 2 (8 years of experience), IQR: Interquartile range, LoA: Bland-Altman limits of agreement, CI: Confidence

CT achieves high diagnostic accuracy in this setting, with an overall sensitivity of 98.7%, specificity of 87.5%, and accuracy of 96.4%. These results are broadly consistent with the pooled estimates from large meta-analytic and multicentre benchmarking data,[1,8] yet provide important granularity regarding the relative diagnostic weight of individual CT signs, the imaging and histopathological characteristics of discordant cases, and the extent of interobserver variability across experience levels.

Overall CT Diagnostic Performance and Comparison with the Literature

The sensitivity of 98.7% achieved in the present series is comparable to and, in some comparisons, exceeds figures reported in seminal large-scale multidetector CT studies. Pickhardt et al.[8] reported a sensitivity of 98.5% and a specificity of 98.0% in a retrospective analysis of 2,871 adults spanning a decade of multidetector CT practice at a single academic centre. The Cochrane systematic review by Rud et al.[1], incorporating 64 studies and 10,280 participants, reported a pooled sensitivity of approximately 95% and specificity of 94% for CT in suspected appendicitis. The slightly lower specificity observed in our series (87.5%) may be partly explained by the proportion of pathology-negative appendectomy specimens within this selected surgical cohort (40/195, 20.5%). Importantly, this figure represents the proportion of patients without histopathological evidence of acute appendicitis among surgically treated patients included in the study, rather than an institutional negative appendectomy rate. Therefore, it should be interpreted in the context of the study design and patient-selection criteria. This discrepancy may be attributable to the tertiary-referral patterns in a training hospital setting, in which higher-risk or diagnostically challenging cases are concentrated, and to a patient-selection effect inherent to centres where CT is routinely employed for all clinically suspected appendicitis cases, irrespective of pre-test probability.

The 2024 IDSA guideline update, representing the most current evidence-based imaging recommendation, conditionally supports CT as the initial imaging modality in non-pregnant adults with suspected acute appendicitis, acknowledging that its accuracy justifies direct CT use without necessitating additional imaging studies.[6] Our data support this recommendation: CT correctly excluded appendicitis in 35 of 37 CT-negative cases, and the two false negatives were both detected clinically through persistent symptom evaluation, independent of imaging.

Individual CT Signs: Relative Diagnostic Weight

Periappendiceal fat stranding emerged as the single most powerful CT predictor of histopathologically confirmed appendicitis in our series (OR: 173.83), present in 96.1% of pathology-positive cases. This finding is consistent with the broader radiology literature, which identifies periappendiceal inflammatory changes as the imaging hallmark of appendiceal inflammation.[9,11] Fat stranding mechanistically reflects exudative oedema and vascular engorgement in the mesoappendix and adjacent fat, and its identification is generally robust across experience levels-as confirmed by the almost perfect interobserver agreement (κ=0.805) in our subsample.

Contrast enhancement demonstrated the highest sensitivity among individual CT features (96.8%) and had a strong independent association (OR: 120.00). Appendiceal wall enhancement reflects intact mural vascularity in the setting of active inflammation and is an earlier and more reliable sign than luminal distension alone.[9,10] Wall thickening, although less specific (specificity 85.0%) than fat stranding, was present in nearly 90% of confirmed cases (OR: 49.23), which makes it a sensitive but not highly discriminatory feature when considered in isolation.

Free fluid reached a high specificity of 95.0% and a PPV of 97.4%, rendering it a powerful confirmatory sign when present, notwithstanding its moderate sensitivity (47.7%). The predominantly periappendiceal distribution of free fluid in our series suggests that when present, it reliably indicates an active periappendiceal inflammatory process rather than coincidental findings such as ovarian follicular fluid or nonspecific pelvic fluid.

The Appendicolith: A Radiologically Identified but Statistically Non-significant Finding

The absence of a statistically significant association between appendicolith and confirmed appendicitis in the present series (OR:

2.04; p=0.151) warrants specific attention. This finding may appear discordant with the rapidly expanding literature on appendicolith-associated appendicitis. However, it should be interpreted in the context of the study design: This analysis assessed whether the presence of an appendicolith predicts histopathologically confirmed appendicitis among patients who have already undergone surgery-not whether it predicts complicated disease or failure of conservative management. Both populations (appendicitis-confirmed and negative appendectomy) harboured appendicoliths at different frequencies (26.5% vs. 15.0%), although sample size limitations prevented the difference from reaching statistical significance.

The clinical relevance of appendicolith detection, however, lies not in confirming the diagnosis of appendicitis per se, but in stratifying disease severity and guiding management decisions in the antibiotic-first era. A large multicentre Finnish study by Sula et al.[14] demonstrated that, among 3,085 patients with CT-confirmed appendicitis, the presence of an appendicolith was associated with a markedly elevated risk of complicated appendicitis (47.1% vs. 21.5%; p<0.001), with larger appendicolith diameter, a base-of-appendix location, and heterogeneous mural enhancement further amplifying this risk. Similarly, Oktay et al.[15] found that CT-detected appendicoliths in paediatric appendicitis were associated with significantly larger appendiceal diameters (10 mm vs. 8 mm; p=0.001), and Weitzner et al.[16] demonstrated that appendicolith presence and size independently predicted CT-histopathological discordance and failure of conservative management. Taken together, the available evidence supports systematic documentation of the presence and characteristics of appendicoliths in CT reports-not as a binary diagnostic sign, but as a prognostically meaningful imaging biomarker that shapes decision-making about non-operative management.

CT-Detected Perforation: The Kappa Paradox and Diagnostic Limitations

CT-detected perforation demonstrated a PPV of 100% in the present series (all 14 CT-reported perforations were pathologically confirmed), yet it failed to reach statistical significance (p=0.078) as a predictor of histopathologically complicated disease and yielded a near-zero kappa for interobserver agreement (κ=-0.034). Both observations are readily explained by well-documented phenomena. The lack of statistical significance reflects the limited sensitivity of CT for detecting perforation, which leads to underestimation of the true histopathological perforation rate; perforation identified on CT represents a subset of cases with macroscopic or extraluminal gas/fluid collections, whereas histopathological assessment detects subtler transmural involvement. The near-zero kappa reflects the well-known kappa paradox: when prevalence is very low, Cohen's kappa may substantially underestimate agreement beyond chance, even when observed concordance is high, as widely discussed in the methodological literature. The 92.5% observed concordance for perforation status in our subsample illustrates this phenomenon clearly.[24,25]

False-positive Cases: Lymphoid Hyperplasia as the Principal Mimic

Lymphoid hyperplasia accounted for 80% (4/5) of false-positive cases in our series; one of these cases also demonstrated concurrent fibrous obliteration. This finding is consistent with the established radiological literature identifying lymphoid hyperplasia as the most important CT mimic of acute appendicitis, particularly in younger patients.[12,13] Lymphoid hyperplasia produces reactive follicular expansion in the lamina propria, leading to appendiceal wall thickening and-critically-periappendiceal inflammatory changes, secondary to reactive enlargement of regional mesenteric lymph nodes and oedematous mesenteric fat. In cases where multiple secondary CT signs are present (as in our two largest-diameter false-positive cases, in which all secondary CT parameters were positive), a false-positive diagnosis is virtually inevitable on CT alone (Figure 3).

Ultrasound-based discrimination between appendicitis and lymphoid hyperplasia has been investigated, with a lamina propria thickness below 1 mm identified as the most effective ultrasound parameter for this differentiation.[12,13] Our findings reinforce the value of the systematic

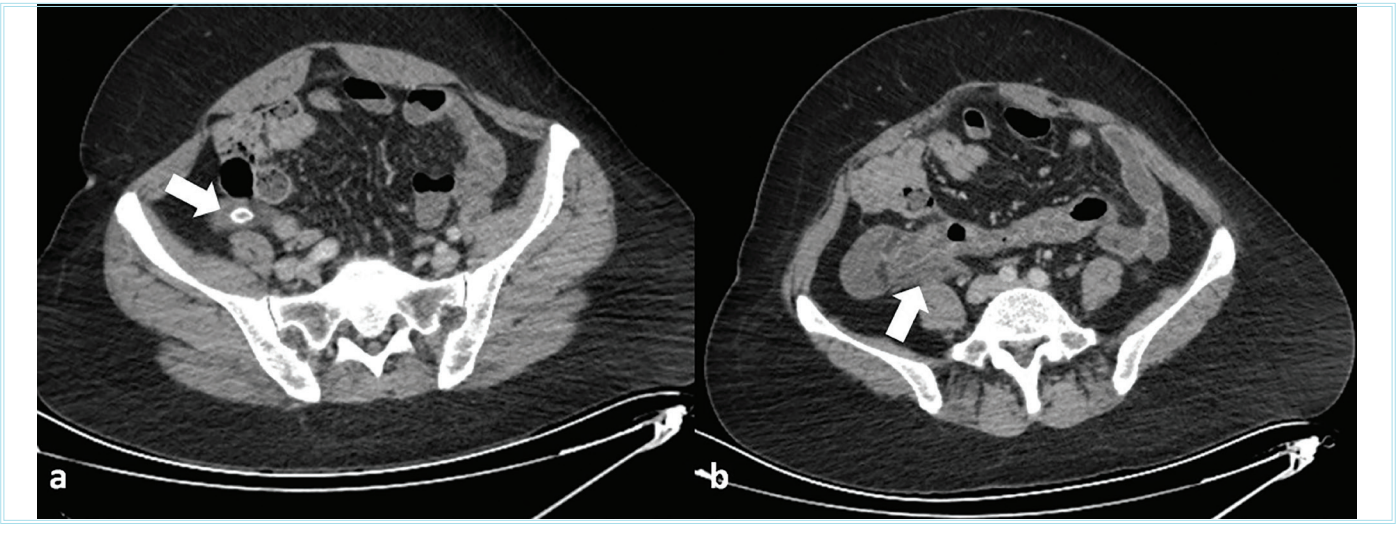


Figure 3. False-positive CT in a 47-year-old female patient with lymphoid hyperplasia mimicking acute appendicitis

(a, b) Axial CT images show an appendicolith in panel (a) and an appendix measuring 13 mm in maximal diameter, with associated periappendiceal inflammatory changes highly suggestive of acute appendicitis. Despite the convincing radiological appearance, histopathological examination demonstrated no evidence of appendicitis and confirmed lymphoid hyperplasia

CT: Computed tomography

integration of clinical pre-test probability with imaging interpretation; in young patients—particularly those aged under 25 years without fever, leukocytosis, or typical pain migration—a degree of diagnostic restraint may be warranted before committing to surgical intervention based solely on CT findings. This argument is also supported by risk-stratification approaches incorporating clinical scores.[3]

FN Cases: Early-stage and Gangrenous Appendicitis

Both FN cases in our series illustrate important and well-described CT diagnostic pitfalls. The first—a 20-year-old male with an appendiceal diameter of 5.4 mm and no secondary signs—represents the classic early-stage appendicitis below the CT detection threshold, in which luminal obstruction has not yet produced sufficient luminal distension or secondary inflammatory changes detectable on cross-sectional imaging (Figure 1). Appendiceal diameters below 6-7 mm in the clinical context of right lower quadrant pain do not reliably exclude early appendicitis; such patients may require close clinical observation, serial examinations, or repeat imaging.[23]

The second case—a 25-year-old woman with an appendiceal diameter of 7.9 mm and gangrenous pathology—exemplifies the paradox of gangrenous appendicitis: transmural necrosis of the appendiceal wall can selectively destroy the mucosal and submucosal layers that mediate the inflammatory and enhancement responses on CT, thereby producing an imaging appearance that underrepresents the severity of the underlying disease (Figure 2). Published case series have demonstrated that gangrenous appendicitis may occasionally present with a poorly enhancing, thin-walled appendix that does not fulfil conventional diagnostic criteria, with the diagnosis ultimately established on clinical and intraoperative grounds.[26] That both FN patients were operated on because of sustained clinical suspicion underscores the principle that a negative or equivocal CT examination does not mandate clinical inaction when other disease indicators remain convincing.

Interobserver Agreement

The perfect interobserver agreement for the overall CT diagnosis of appendicitis (κ=1.000) in our subsample of 40 patients is an encouraging finding, suggesting that even with one year of experience, a trainee radiologist can make accurate categorical CT diagnoses of appendicitis when evaluating a properly performed contrast-enhanced examination in a blinded manner. This result aligns with data from Hof et al.[20], who demonstrated that the specificity of CT for diagnosing appendicitis was consistent across radiologists with lower and intermediate experience (94% in both groups), whereas sensitivity improved with greater experience.

At the level of individual CT features, a consistent pattern emerges: objective, well-defined morphological features (appendicolith: κ=0.805; periappendiceal fat stranding: κ=0.805) achieve higher agreement than subjective or prevalence-dependent features (free fluid: κ=0.504; CT-detected perforation: κ=-0.034, reflecting the kappa paradox). This hierarchy has important implications for structured CT reporting: standardised reporting templates that explicitly require a binary assessment of each primary inflammatory sign—rather than a gestalt impression—may reduce the experience-dependent variability observed in secondary features, such as free fluid.

The excellent reproducibility of appendiceal diameter measurement (ICC=0.994; Bland-Altman limits of agreement -0.79 to +0.76 mm) is clinically relevant, particularly given the emerging evidence supporting appendiceal diameter as a prognostic biomarker for complicated disease and failure of non-operative management. Near-perfect diameter reproducibility across experience levels indicates that this metric can be reliably incorporated into clinical decision algorithms and risk-stratification tools without concern about measurement variability.

Sex-specific Differences

A statistically significant sex-based difference in histopathological confirmation rates was observed: 87.5% of male patients had confirmed appendicitis, compared with 71.7% of female patients (p=0.011). This differential likely reflects the greater diagnostic challenge posed by fertile women, in whom gynaecological conditions such as ovarian cysts, adnexitis, and ectopic pregnancy can closely mimic appendicitis both clinically and on CT. The higher proportion of pathology-negative appendectomy specimens among women is a well-established observation in the appendicitis literature and underscores the diagnostic challenge posed by gynaecological mimics in women of reproductive age.[6,7]

Pathology-negative Appendectomy Specimens and Clinical Implications

In this selected cohort of patients who underwent appendectomy after preoperative CT for clinically suspected acute appendicitis, 40 of 195 patients (20.5%) had no histopathological evidence of acute appendicitis. This proportion should be interpreted as the pathology-negative subgroup within the study population, rather than as a direct institutional quality metric or a general negative appendectomy rate. Several factors may have contributed to the presence of pathology-negative cases, including the training hospital setting, a heterogeneous referral base, the predominance of young patients in whom lymphoid hyperplasia may mimic acute appendicitis, and the tendency to proceed to surgery when clinical suspicion persists despite diagnostic uncertainty. Structured risk-stratification pathways incorporating clinical scoring systems such as the Appendicitis Inflammatory Response (AIR) score may help refine decision-making and should be prospectively evaluated at our institution.[27,28]

Study Limitations

Several limitations of this study should be acknowledged. First, the retrospective single-centre design may limit the generalisability of the findings to other clinical settings. Second, the relatively small numbers of complicated appendicitis cases (n=18) and discordant cases (n=7) limited the statistical power of subgroup analyses. Third, clinical scoring data, including Alvarado and AIR scores, were not systematically available for the entire cohort, precluding formal evaluation of the incremental diagnostic value of CT beyond clinical assessment. Finally, although a consensus CT diagnosis served as the primary CT outcome, the consensus process itself may have introduced bias in borderline cases.

Conclusion

This single-centre, retrospective study confirms that contrast-enhanced CT has high diagnostic performance for acute appendicitis in an adult surgical population, with a sensitivity of 98.7%, specificity of 87.5%, and overall accuracy of 96.4%. Periappendiceal fat stranding represents the strongest individual CT predictor (OR: 173.83), while a CT-detected appendicolith does not independently reach diagnostic significance for confirming appendicitis, although its prognostic value for complicated disease is well supported by the wider literature. Gangrenous appendicitis

with an attenuated inflammatory response on CT and early-stage disease below the luminal threshold are the principal sources of FN diagnoses. Interobserver agreement for the overall CT diagnosis is perfect across experience levels, whereas individual secondary signs—particularly free fluid and perforation assessment—demonstrate experience-dependent variability. These findings support the current evidence-based recommendation for CT as the primary diagnostic imaging modality in non-pregnant adult patients with suspected acute appendicitis and highlight the continued importance of integrated clinical assessment for CT-negative cases with persistent symptomatology.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Kastamonu University Non-Interventional Clinical Research Ethics Committee (approval number: 2026-49, date: 19.03.2026).

Informed Consent: Retrospective Study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: N.S., Y.A.Y., Concept: E.Ö., Design: E.Ö., Data Collection or Processing: E.Ö., A.C., N.S., Y.A.Y., Analysis or Interpretation: E.Ö., A.C., Literature Search: E.Ö., A.C., N.S., Y.A.Y., E.Ö., Writing: E.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Rud B, Vejborg TS, Rappeport ED, Reitsma JB, Wille-Jørgensen P. Computed tomography for diagnosis of acute appendicitis in adults. Cochrane Database Syst Rev. 2019;2019:CD009977.
2. Börner N, Kappenberger AS, Weber S, Scholz F, Kazmierczak P, Werner J. The acute abdomen: structured diagnosis and treatment. Dtsch Arztebl Int. 2025;122:137-44.
3. Deboni VS, Rosa MI, Lima AC, Graciano AJ, Garcia CE. The appendicitis inflammatory response score for acute appendicitis: is it important for early diagnosis? Arq Bras Cir Dig. 2022;35:e1686.
4. Chen KC, Arad A, Chen KC, Storrar J, Christy AG. The clinical value of pathology tests and imaging study in the diagnosis of acute appendicitis. Postgrad Med J. 2016;92:611-9.
5. Andre JB, Sebastian VA, Ruchman RM, Saad SA. CT and appendicitis: evaluation of correlation between CT diagnosis and pathological diagnosis. Postgrad Med J. 2008;84:321-4.
6. Bonomo RA, Tamma PD, Abrahamian FM, et al. 2024 Clinical practice guideline update by the Infectious Diseases Society of America on complicated intra-abdominal infections: diagnostic imaging of suspected acute appendicitis in adults, children, and pregnant people. Clin Infect Dis. 2024;79(Suppl 3):S94-103.
7. Di Saverio S, Podda M, De Simone B, et al. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. World J Emerg Surg. 2020;15:27.
8. Pickhardt PJ, Lawrence EM, Pooler BD, Bruce RJ. Diagnostic performance of multidetector computed tomography for suspected acute appendicitis. Ann Intern Med. 2011;154:789-96.
9. Aydın S, Karavas E, Şenbil DC. Imaging of acute appendicitis: advances. World J Gastrointest Surg. 2022;14:370-3.
10. Gurian MS, Kovanlikaya A, Beneck D, Baron KT, John M, Brill PW. Radiologic-pathologic correlation in acute appendicitis: can we use it as a quality measure to assess interpretive accuracy of radiologists? Clin Imaging. 2011;35:421-3.
11. Simianu VV, Shamitoff A, Hippe DS, et al. The reliability of a standardized reporting system for the diagnosis of appendicitis. Curr Probl Diagn Radiol. 2017;46:267-74.
12. Aydin S, Tek C, Ergun E, Kazci O, Kosar PN. Acute appendicitis or lymphoid hyperplasia: how to distinguish more safely? Can Assoc Radiol J. 2019;70:354-60.
13. Tanabe M, Maeda K, Kuninaka H, et al. An infant autopsy case of acute appendicitis with lymphoid hyperplasia. Pediatr Rep. 2025;17:96.
14. Sula S, Paananen T, Tammilehto V, et al. Impact of an appendicolith and its characteristics on the severity of acute appendicitis. BJS Open. 2024;8:zrae093.
15. Oktay C, Goksu M, Yavuz S. Prevalence of appendicolith in children with acute appendicitis and its correlation with disease severity. North Clin Istanb. 2023;10:631-5.
16. Weitzner ZN, Chung A, Naini BV, Graham D, Livingston EH. Correlation of computed tomography, pathological findings, and clinical outcomes for appendicoliths in appendicitis. Ann Surg Open. 2023;4:e280.
17. Zhang D, Wang S, Li H, et al. Retrospective analysis of 331 acute appendicitis patients: how appendicolith and CT features aid in differentiating complicated vs. uncomplicated appendicitis. BMC Gastroenterol. 2025;26:40.
18. Coutureau J, Millet I, Taourel P. CT of acute abdomen in the elderly. Insights Imaging. 2025;16:95.
19. Sula S, Kujala M, Tammilehto V, et al. Prognostic CT-imaging findings for complicated acute appendicitis: a prospective cohort study. Scand J Surg. 2026;115:42-9.
20. in't Hof KH, Krestin GP, Steijerberg EW, et al. Interobserver variability in CT scan interpretation for suspected acute appendicitis. Emerg Med J. 2009;26:92-4.
21. Alsayaf Alghamdi AG, Alzhrani SM, Fayraq A, Alzahrani SA. Predictive value of clinical and CT scan findings for complicated appendicitis: a retrospective analysis. Cureus. 2025;17:e88948.
22. Issaiy M, Zarei D, Saghazadeh A. Artificial intelligence and acute appendicitis: a systematic review of diagnostic and prognostic models. World J Emerg Surg. 2023;18:59.
23. Moris D, Paulson EK, Pappas TN. Diagnosis and management of acute appendicitis in adults: a review. JAMA. 2021;326:2299-311.
24. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43:543-9.
25. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46:423-9.
26. Suzuki T, Matsumoto A, Sugiki D, Akao T, Matsumoto H. Clinical prediction model for gangrenous appendicitis: a retrospective single-center study. Scand J Surg. 2025;114:210-7.
27. Von-Mühlen B, Franzon O, Beduschi MG, Kruel N, Lupselo D. AIR score assessment for acute appendicitis. Arq Bras Cir Dig. 2015;28:171-3.
28. Baştürk T, Duran M, Baştürk S. Evaluation of computed tomography (CT) appendicitis score and laboratory parameters in acute appendicitis with and without CT-detected appendicolith. TJTES. 2025;31:651-60.