



Diagnostic performance and inter-observer variability to differentiate between T1- and T2-stage gallbladder cancers using multi-detector row CT

Yong Jae Kwon¹ · Kyoung Doo Song¹ · Seong Eun Ko¹ · Jeong Ah Hwang¹ · Minji Kim²

Received: 9 December 2021 / Revised: 7 February 2022 / Accepted: 9 February 2022 / Published online: 22 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Purpose To evaluate the diagnostic performance and inter-observer variability of differentiating T1 and T2 gallbladder (GB) cancers using multi-detector row CT (MDCT).

Methods This retrospective study included 151 patients with surgically confirmed T1 ($n=49$)- or T2 ($n=102$)-stage GB cancer who underwent contrast-enhanced MDCT from 2016 to 2020. Five radiologists (two experienced and three less experienced) evaluated the T-stage with a confidence level calculated using a six-point scale. GB cancers were morphologically classified into three types: polypoid, polypoid with wall thickening, and wall thickening. The diagnostic performance of T-staging was assessed using receiver operating characteristic (ROC) curve analysis. Sensitivity, specificity, accuracy, positive predictive value, and negative predictive value were calculated based on a binary scale (T1 = positive). Inter-observer agreement was assessed using Fleiss κ statistics.

Results The area under the receiver operating characteristic (ROC) curve of each reviewer for T-staging ranged from 0.69 to 0.80 (median 0.77). The overall accuracy of the five radiologists was 78% (95% confidence interval [CI] 71–84%). Sensitivity was higher and specificity was lower in experienced radiologists than in less experienced radiologists ($P < 0.001$). The overall inter-observer agreement was fair ($\kappa = 0.36$; 95% CI 0.31, 0.41). The overall accuracy for T-stage was 63% (95% CI 48–76), 78% (95% CI 63–88), and 87% (95% CI 77–93) for polypoid, polypoid with wall thickening, and wall thickening type, respectively.

Conclusion The accuracy of MDCT for differentiating T1 and T2 GB cancer is limited, and there is considerable inter-observer variability.

Keywords Gallbladder cancer · Early stage · Computed tomography · Diagnostic performance · Inter-observer variability

Abbreviations

GB	Gallbladder
CT	Computed tomography
MDCT	Multi-detector row computed tomography
MRP	Multiplanar reformation
ROC	Receiver operating characteristic
AUC	Area under the curve
GEE	Generalized estimating equations

Introduction

Gallbladder (GB) cancer is the most common malignant tumor of the gallbladder. Although GB cancer is a relatively rare neoplasm, some countries in Asia, Latin America, and Eastern Europe show a high incidence rate [1, 2]. GB cancer is often diagnosed at an advanced stage, because it presents as asymptomatic or nonspecific symptoms in the early stages and lacks a submucosal layer. However, the use of ultrasonography and computed tomography (CT) has increased the detection of early-stage GB cancer [3].

Complete resection with a negative margin is the only curative treatment for early-stage GB cancer [4]. The extent of surgical resection is usually determined according to the T-stage. Simple cholecystectomy is sufficient for the treatment of T1a GB cancer, and extended cholecystectomy with lymph node dissection is the standard treatment

✉ Kyoung Doo Song
kd3893.song@samsung.com

¹ Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-Ro, Gangnam-Gu, Seoul 06351, Republic of Korea

² Statistics and Data Center, Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea

for T2 GB cancer. However, controversy regarding the optimal surgery for T1b GB cancer remains [4, 5]. The National Comprehensive Cancer Network guidelines and the European Society for Medical Oncology guidelines recommend cholecystectomy with liver resection (wedge resection or segment IVb and V) for T1b-stage GB cancer [6, 7]. However, several studies have shown comparable survival outcomes in patients with T1b GB cancer between simple and extended cholecystectomy [8–13]. In institutions and countries that perform simple cholecystectomy as a treatment for T1b GB cancer, precise preoperative differentiation between T1- and T2-stage GB cancers can provide surgeons with an appropriate surgical plan to decrease the risk of unnecessary over-resection or reoperation rates.

Several studies have investigated the CT criteria and diagnostic accuracy for the T-stage of GB cancer [14–16]. However, these studies were intended to evaluate the diagnostic performance of all T-stage GB cancers, and the study population of each study was not sufficiently large. In addition, there are discordant results in diagnostic performance between these studies, and only one study evaluated the inter-observer variability.

Therefore, the purpose of our study was to evaluate the diagnostic performance and inter-observer variability of differentiating between T1- and T2-stage GB cancers using multi-detector row CT (MDCT).

Materials and methods

This retrospective, single-center study was conducted at a tertiary referral academic center. The Institutional Review Board approved this study and waived the need for written informed consent. The study was conducted in accordance with the Declaration of Helsinki.

Patients

A total of 218 consecutive patients who had undergone surgery for GB cancer between January 2016 and September 2020 were identified through a search of our institution's surgical database. Among them, 162 patients were

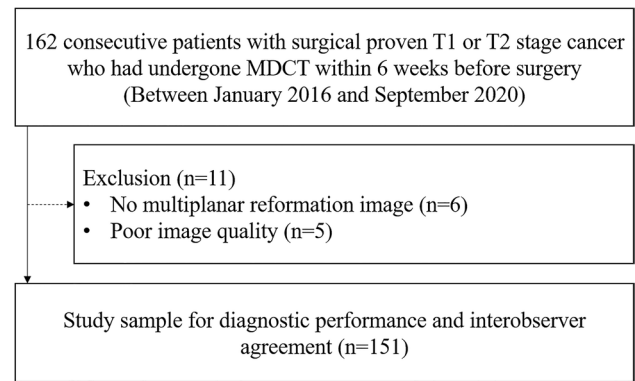


Fig. 1 Inclusion flowchart

included according to the following inclusion criteria: (a) patients with T1 or T2 GB cancers and (b) patients who had undergone MDCT within 6 weeks before surgery. Eleven patients were excluded from the study due to the absence of multiplanar reformation (MRP) images ($n=6$) or poor image quality ($n=5$). A total of 151 patients (81 women and 70 men; mean age, $66.6 \text{ years} \pm 10.4$) were included in the study (Fig. 1).

Image acquisition

Various MDCT scanners were used: Lightspeed VCT, Discovery CT750 HD, and Revolution Frontier of GE Healthcare (Milwaukee, WI, USA), Somatom Definition Flash and Somatom Force of Siemens Healthcare (Erlangen, Germany), and Aquilion and Aquilion One of Canon Medical Systems (Ottawa, Japan). The CT protocol for the evaluation of biliary disease consists of three phases (unenhanced, arterial, and portal venous phases). Arterial and portal phases were obtained 45 and 70 s after the injection of iodinated contrast agent (typically 120 mL of 300 mgI/mL at 4 mL/s). The scanning parameters were as follows: tube voltage, 120 kV; tube current, 120–350 mA; and pitch factor, 0.984–1.375. Automated tube current modulation was used. The slice thickness for axial images was 3 or 3.75 mm, and the slice thickness for coronal and sagittal images was 3 mm. Scan coverage was from the right liver dome to the third portion of the duodenum for unenhanced and arterial phases and from the xiphoid process to the pubic symphysis.

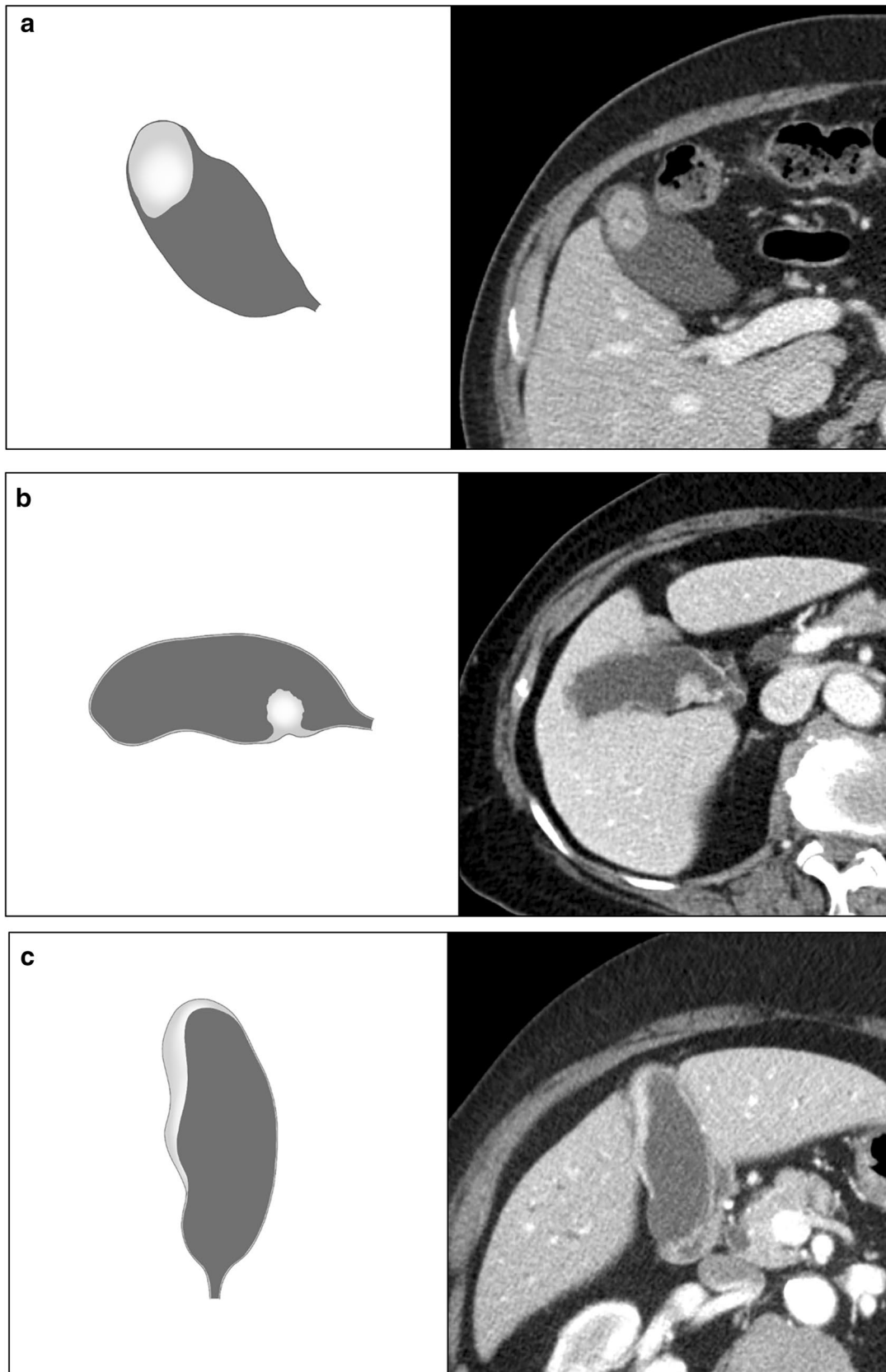


Fig. 2 Tumor morphologic type. **a** Polypoid type, **b** polypoid with wall thickening type. **c** Wall thickening type

Coronal and sagittal MPR images of portal phase images were subsequently generated.

Image review

CT image review was independently performed by five radiologists (two experienced radiologists, with 10 and 11 years of experience in abdominal imaging, and three less experienced radiologists, with 2–6 years of experience in abdominal imaging). Before reviewing the CT images of patients enrolled in this study, previous studies on the staging of gallbladder cancer were shared and reviewed [14, 15]. The reviewers used the criteria suggested by Kim et al. to differentiate between T1 and T2 GB cancers, which are as follows: T1, polypoid lesions without focal thickening of the gallbladder wall, nodular or flat lesions with mucosal enhancement or focal thickening of the inner enhancing layer of the gallbladder wall with clear low-attenuated outer wall; T2, nodular or sessile lesions associated with focal thickening of the gallbladder wall at what was considered to be attachment sites, diffuse wall thickening with heterogeneous enhancement, diffuse wall thickening with strong thick inner wall enhancement and weak enhancement of the outer layer, focal wall thickening with outer surface dimpling at the tumor base [15]. As a training session, CT images and information regarding the T-stage of 40 patients (20 T1 cancers and 20 T2 cancers) who were not included in this study were reviewed by all reviewers.

In the reading session, the reviewers were informed of the purpose of this study. However, the reviewers were blinded to the clinical information, including the staging. The reviewers evaluated the CT images using a picture archiving and communication system (PACS; Centricity Radiology RA 1000; GE Healthcare, Chicago, IL, USA). The reviewers independently evaluated the T-stage of gallbladder cancers using a six-point scale as follows: 1, definitely T1; 2, probably T1; 3, T1 more likely than T2; 4, T2 more likely than T1; 5, probably T2; 6, definitely T2. In separate sessions, two out of five reviewers classified the radiologic morphology of the tumors into the following three categories independently, as either: polypoid, polypoid with focal wall thickening, or wall thickening type (Fig. 2). After independent image review

of the tumor morphology, a consensus was reached if there were discordant results between reviewers.

Clinical data acquisition

The following information was obtained by reviewing the electronic medical records of our institution: age, sex, carbohydrate antigen 19-9 (CA19-9), type of surgery, and pathologic T-stage.

Statistical analysis

Under the assumption that the proportion of accuracy under the null hypothesis (P_0) is 0.85, a sample size of 148 achieved 80% power to detect a difference of -0.10 using a two-sided Z-test with a significance level (α) of 0.050.

The diagnostic performance of each reviewer using the six-point scale was assessed using receiver operating characteristic (ROC) curve analysis, and the area under the ROC (AUC) was calculated. In addition, we calculated the sensitivity, specificity, accuracy, positive predictive value, and negative predictive value based on a binary scale. 1, 2, and 3 of the six-point scale were assumed to be T1, and 4, 5, and 6 were assumed to be T2. The T1-stage was defined as positive. To obtain the overall diagnostic performance of all radiologists, staging was determined according to the results of a majority of the reviewers. Inter-observer agreement for T-staging based on the binary scale of T-stage was assessed using Fleiss κ statistics. Inter-observer agreement between the two reviewers for the morphologic tumor type was evaluated using κ statistics. Kappa value interpretations were as follows: $\kappa \leq 0.2$, poor; $0.2 < \kappa \leq 0.4$, fair; $0.4 < \kappa \leq 0.6$, moderate; $0.6 < \kappa \leq 0.8$, substantial; $0.8 < \kappa \leq 1.0$, excellent. The inter-observer agreement for T-staging was compared between radiologists' experience using two independent sample z tests on the basis of asymptotic normality. The diagnostic performance (sensitivity, specificity, and accuracy) for T-staging was compared between radiologists' experiences by analysis using generalized estimating equations (GEE). A two-sided P-value less than 0.05 was considered to be statistically significant. Statistical analysis was performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

Table 1 Demographics of Enrolled Patients (n = 151)

Age (year)*	67 ± 10 (38–89)
Male	70 (46)
Female	81 (54)
T-stage	
T1a	37 (25)
T1b	12 (8)
T2	102 (68)
CA19-9 (U/mL)*	59.6 ± 334
Type of surgery	
Simple cholecystectomy	58(38)
Radical resection	80(53)
Radical resection after simple cholecystectomy	13(9)

Unless otherwise indicated, data are numbers of patients with percentages in parenthesis

*Data are means ± standard deviations

Results

The baseline characteristics of the patients are summarized in Table 1. Our study population consisted of 49 (32%) patients with T1 GB cancers and 102 (68%) patients with T2 GB cancers.

Diagnostic performance

The AUC, sensitivity, specificity, accuracy, PPV, and NPV of each reviewer for T-staging are shown in Table 2 and Fig. 3. The AUC ranged 0.69–0.80. The overall accuracy of the five radiologists was 78%. The sensitivity was higher for experienced reviewers ($P < 0.001$), while specificity was higher in less experienced reviewers ($P < 0.001$).

Inter-observer agreement for T-staging

The results of the inter-observer agreement for T-staging are shown in Table 3. For the assessment of T-stage, the inter-observer agreement of all reviewers showed fair reliability ($\kappa = 0.36$; 95% confidence interval: 0.31, 0.41). Although inter-observer agreement of the more experienced reviewers was higher than that of the less experienced reviewers, the difference was not statistically significant ($\kappa = 0.54$ vs 0.38, $P = 0.093$).

Analysis according to morphologic tumor type

According to the radiologic morphology, 41 (27.2%), 40 (26.5%), and 70 (46.4%) tumors were classified as polypoid,

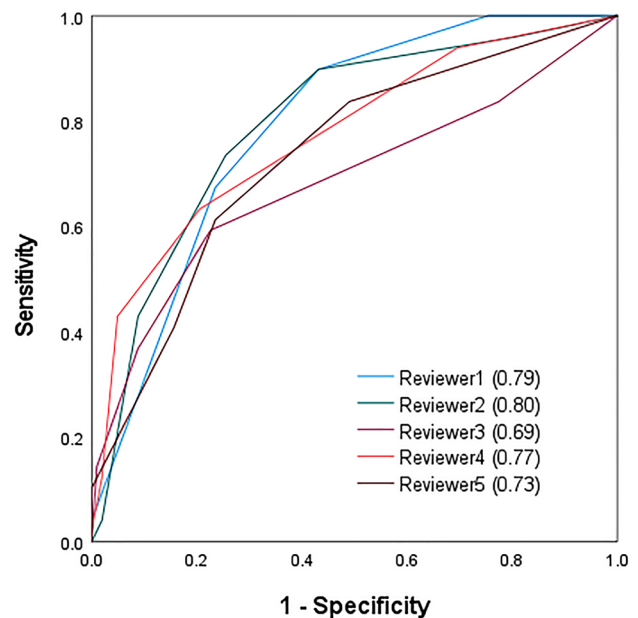


Fig. 3 ROC curve analysis for T-staging

Table 2 Diagnostic performance in T-staging of T1 and T2 gallbladder cancers

Reviewer	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
Reviewer 1	0.79 (0.72, 0.86)	67 (53, 79)	76 (67, 84)	74 (66, 80)	58 (45, 70)	83 (74, 89)
Reviewer 2	0.80 (0.72, 0.87)	73 (60, 84)	75 (65, 82)	74 (67, 80)	58 (46, 70)	85 (77, 91)
Reviewer 3	0.69 (0.59, 0.78)	37 (25, 51)	91 (84, 95)	74 (66, 80)	67 (48, 81)	75 (67, 82)
Reviewer 4	0.77 (0.69, 0.85)	43 (30, 57)	95 (89, 98)	78 (71, 84)	81 (62, 91)	78 (70, 84)
Reviewer 5	0.73 (0.65, 0.82)	41 (28, 55)	84 (76, 90)	70 (62, 77)	56 (40, 70)	75 (66, 82)
Overall		59 (45, 72)	87 (79, 92)	78 (71, 84)	69 (54, 81)	82 (73, 88)

Data in parentheses are 95% confidence intervals, AUC = area under the receiver operating characteristic curve, NPV negative predictive value, PPV positive predictive value

AUC was assessed using a six-point scale for T-staging of gallbladder cancer: 1 definitely T1, 2 probably T1, 3 T1 more likely than T2, 4 T2 more likely than T1, 5 probably T2, 6 definitely T2. Sensitivity, specificity, accuracy, PPV and NPV were assessed by using a binary scale, as follows: T1 (positive) vs T2 (negative)

Table 3 Results of Inter-observer Agreement for T-staging

	All reviewers (<i>n</i> =5)	Experienced reviewers (<i>n</i> =2)	Less experienced reviewers (<i>n</i> =3)	<i>P</i> value*
All tumors	0.36 (0.31, 0.41)	0.54 (0.38, 0.70)	0.38 (0.29, 0.48)	0.093
Morphologic tumor type				
Polypoid	0.07 (− 0.03, 0.16)	0.02 (− 0.28, 0.33)	0.17 (− 0.01, 0.35)	0.414
Polypoid + wall thickening	0.25 (0.16, 0.35)	0.33 (0.02, 0.64)	0.43 (0.26, 0.61)	0.564
Wall thickening	0.32 (0.25, 0.40)	0.43 (0.20, 0.67)	0.39 (0.26, 0.53)	0.758

Inter-observer agreement based on the binary scale was assessed by using Fleiss κ statistics

**P* value for comparison of Fleiss κ value between experienced reviewers and less experienced reviewers. *P* values were calculated by two independent sample *z* test on the basis of asymptotic normality. Data in parentheses are 95% confidence intervals

polypoid with wall thickening, and wall thickening type, respectively. The inter-observer agreement of the two reviewers for the classification of morphologic tumor type was substantial ($\kappa=0.70$).

The overall accuracies for the T-stage were 63%, 78%, and 87% for polypoid, polypoid with wall thickening, and wall thickening type, respectively. In the polypoid type, the sensitivity and accuracy were higher ($P<0.001$ and $P=0.002$, respectively), and the specificity was lower in experienced reviewers ($P=0.009$) (Supplementary Figure S1–S3). In polypoid with wall thickening type, the sensitivity showed no difference ($P=0.823$), but the specificity and accuracy were lower for experienced reviewers ($P<0.001$ and $P=0.012$, respectively). In the wall thickening type, the sensitivity, specificity, and accuracy were not different between experienced reviewers and less experienced reviewers ($P=0.141$, $P=0.062$, and $P=0.252$, respectively). The kappa values for the T-stage of the five reviewers were 0.07 (− 0.03, 0.16), 0.25 (0.16, 0.36), 0.32 (0.25, 0.40) for polypoid, polypoid with wall thickening, and wall thickening type, respectively (Table 3).

Analysis of incorrect T-staging cases and poor inter-observer agreement cases

Seven patients with T1 GB cancers were incorrectly staged as T2 by all five reviewers. Four patients had polypoid wall thickening-type GB cancers, and three patients had GB cancers with wall thickening (Fig. 4). There were no patients in whom T2 GB cancers were incorrectly staged as T1 by all five reviewers (Fig. 5).

There were 43 patients with poor inter-observer agreement (this refers to a patient in whom the results of five reviewers for T-staging were divided by 2:3). Among 43 patients, 24, 9, and 10 patients were polypoid, polypoid with wall thickening, and wall thickening type, respectively. Poor inter-observer agreement was seen in 59% (24/41) of patients with polypoid type, 23% (9/40) of patients with polypoid with wall thickening, and 14% (10/70) of patients with wall thickening type, and the proportion of patients with poor inter-observer agreement was different according to the morphologic type ($P<0.001$) (Table 4).

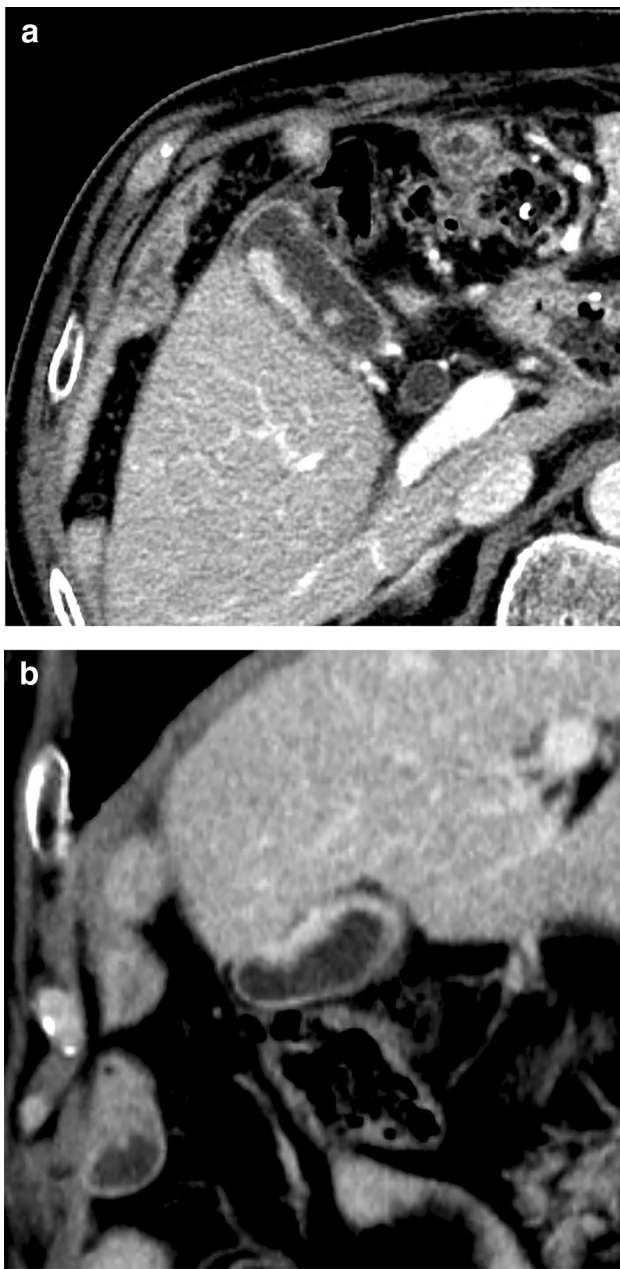


Fig. 4 A case with good inter-observer agreement between reviewers but inaccurate T-staging. The case was a 71-year-old man with GB cancer, with a tumor of the wall thickening type. All reviewers staged the tumor as T2, but the pathologic results were T1a. Axial **a** and coronal, **b** images

Discussion

Recent studies have demonstrated that simple cholecystectomy has similar recurrence and survival rates to radical cholecystectomy for T1b GB cancer [8]. Based on the results of these studies, it is important to distinguish T1- and T2-stage GB cancers preoperatively. Our study showed that the diagnostic performance of radiologists for differentiating T1 and T2 GB cancers using MDCT was a median AUC of 0.77 an overall accuracy of 78%, and the inter-observer agreement between radiologists was fair ($\kappa=0.36$). In our results, the diagnostic performance and inter-observer agreement were not satisfactory.

In previous studies, the diagnostic performance for T-staging using CT has been reported. Yoshimitsu et al. reported an overall accuracy for T1-stage GB cancer of 86%, while Kim et al. reported 94% [15, 16]. Compared to the results of previous studies, the overall accuracy for T-staging was low at 78% in our study. Even the radiologist with the highest accuracy had an accuracy of only 80%. The diagnostic performance for T-staging can be influenced by the experience level of the radiologists and the characteristics of the tumors included. To minimize the effect of these factors on the results, radiologists with various experience levels were included in our study. In addition, we enrolled a sufficient number of patients to minimize the random error resulting from the small sample size.

Diagnostic performance was affected by the experience level of radiologists. Experienced radiologists showed slightly higher AUCs for T-staging than less experienced radiologists (0.79 and 0.80 vs 0.69, 0.77, 0.73, respectively). However, the accuracy was not different between experienced and less experienced radiologists (74% vs. 74% vs. 70%, 74%, and 78%, respectively). Interestingly, the sensitivity was higher, but the specificity was lower for experienced radiologists. In other words, experienced radiologists tended to understage, and less experienced radiologists tended to overstage. In addition, the inter-observer agreement of the experienced reviewers was higher than that of the less experienced reviewers, although the difference was not statistically significant.

Fig. 5 Cases of poor inter-observer agreement for T-staging. **a** A 70-year-old man with polypoid-type T1a GB cancer. Three reviewers staged T1 and two reviewers staged T2. Axial image. **b** 64-year-old man with polypoid-type T2 GB cancer. Three reviewers staged T2, and two reviewers staged T1. Axial image. **c** A 61-year-old man with polypoid wall thickening-type T1a GB cancer. Three reviewers staged T1 and two reviewers staged T2. Coronal image. **d** A 50-year-old woman with polypoid wall thickening-type T2 GB cancer. Three reviewers staged T2, and two reviewers staged T1. Coronal image

In our study, all tumors were classified into three morphologic types: polypoid, polypoid with wall thickening, and wall thickening type. The overall accuracy was low in the polypoid type (63%, 78%, and 87% for polypoid, polypoid with wall thickening, and wall thickening type, respectively), as was the inter-observer agreement (0.07, 0.25, 0.32, for polypoid, polypoid with wall thickening, and wall thickening type, respectively). For polypoid-type tumors with a broad base, it is difficult to distinguish T1 and T2 because the boundary between the tumor and the gallbladder wall can be ambiguous. This may be one of the reasons for the low accuracy and low inter-observer agreement in polypoid-type tumors.

Our study has several limitations. First, this study was conducted at a single tertiary hospital. Therefore, our results may not be generalizable to other hospitals in different countries. Second, we included only T1- and T2-stage GB cancers. Therefore, in real clinical situations where everything from benign lesions to advanced GB cancers exists, the results of diagnostic accuracy and inter-observer variability may vary.

In conclusion, the accuracy of MDCT for differentiating T1 and T2 GB cancers is limited, and there is considerable inter-observer variability. The diagnostic accuracy and inter-observer variability were worst for polypoid lesions. There is a need for further research on how to improve the accuracy of preoperative T-staging in early GB cancer.



Table 4 Diagnostic Performance in T-Staging of T1 and T2 Gallbladder Cancer according to morphologic type

Morphologic type	Reviewer	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
Polypoid <i>n</i> = 41 (34/7) ^b	Reviewer 1	0.61 (0.42, 0.79)	82 (66, 92)	NA ^a	68 (53, 80)	80 (64, 90)	NA ^a
	Reviewer 2	0.55 (0.34, 0.77)	85 (70, 94)	14 (3, 51)	73 (58, 84)	83 (67, 92)	17 (3, 56)
	Reviewer 3	0.67 (0.41, 0.93)	35 (21, 52)	71 (36, 92)	41 (28, 57)	86 (60, 96)	19 (8, 37)
	Reviewer 4	0.78 (0.63, 0.92)	47 (31, 63)	100 (65, 100)	56 (41, 70)	100 (81, 100)	28 (14, 48)
	Reviewer 5	0.60 (0.44, 0.76)	53 (37, 69)	29 (8, 64)	49 (34, 64)	78 (58, 90)	11 (3, 33)
	Overall			71 (54, 83)	29 (8, 64)	63 (48, 76)	83 (65, 92)
Polypoid + wall thickening <i>n</i> = 40 (9/31) ^b	Reviewer 1	0.64 (0.47, 0.81)	22 (6, 55)	71 (53, 84)	60 (45, 74)	18 (5, 48)	76 (58, 88)
	Reviewer 2	0.64 (0.42, 0.86)	44 (19, 73)	74 (57, 86)	68 (52, 80)	33 (14, 61)	82 (64, 92)
	Reviewer 3	0.61 (0.36, 0.87)	44 (19, 73)	90 (75, 97)	80 (65, 90)	57 (25, 84)	85 (69, 93)
	Reviewer 4	0.73 (0.53, 0.93)	33 (12, 65)	97 (84, 99)	83 (68, 91)	75 (30, 95)	83 (68, 92)
	Reviewer 5	0.49 (0.27, 0.7)	11 (2, 44)	94 (79, 98)	75 (60, 86)	33 (6, 79)	78 (63, 89)
	Overall			33 (12, 65)	90 (75, 97)	78 (63, 88)	50 (19, 81)
Wall thickening <i>n</i> = 70 (6/64) ^b	Reviewer 1	0.71 (0.5, 0.91)	50 (19, 81)	88 (77, 94)	84 (74, 91)	27 (10, 57)	95 (86, 98)
	Reviewer 2	0.67 (0.41, 0.92)	50 (19, 81)	81 (70, 89)	79 (68, 87)	20 (7, 45)	95 (85, 98)
	Reviewer 3	0.44 (0.09, 0.8)	33 (10, 70)	94 (85, 98)	89 (79, 94)	33 (10, 70)	94 (85, 98)
	Reviewer 4	0.69 (0.51, 0.88)	33 (10, 70)	94 (85, 98)	89 (79, 94)	33 (10, 70)	94 (85, 98)
	Reviewer 5	0.68 (0.43, 0.94)	17 (3, 56)	86 (75, 92)	80 (69, 88)	10 (2, 40)	92 (82, 96)
	Overall			33 (10, 70)	92 (83, 97)	87 (77, 93)	29 (8, 64)

Data in parentheses are 95% confidence intervals, *AUC* area under the receiver operating characteristic curve, *NPV* negative predictive value, *PPV* positive predictive value

AUC was assessed using a six-point scale for T-staging of gallbladder cancer: 1 definitely T1, 2 probably T1, 3 T1 more likely than T2, 4 T2 more likely than T1, 5 probably T2, 6 definitely T2

Sensitivity, specificity, accuracy, PPV and NPV were assessed using a binary scale as follows: T1 (positive) vs T2 (negative)

^aIn polypoid tumors, the value was not estimated because the reviewer 1 did not stage all T2 tumors as T2

^bData in parentheses are number of T1 cancers/number of T2 cancers

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00261-022-03450-3>.

Declarations

Conflict of interest The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

References

- Randi G, Franceschi S, La Vecchia C (2006) Gallbladder cancer worldwide: geographical distribution and risk factors. *Int J Cancer* 118:1591–1602. <https://doi.org/10.1002/ijc.21683>
- Torre LA, Siegel RL, Islami F, Bray F, Jemal A (2018) Worldwide Burden of and Trends in Mortality From Gallbladder and Other Biliary Tract Cancers. *Clin Gastroenterol Hepatol* 16:427–437. <https://doi.org/10.1016/j.cgh.2017.08.017>
- Kalra N, Suri S, Gupta R, Natarajan SK, Khandelwal N, Wig JD, Joshi K (2006) MDCT in the staging of gallbladder carcinoma. *AJR Am J Roentgenol* 186:758–762. <https://doi.org/10.2214/ajr.04.1342>
- Aloia TA, Járufe N, Javle M, Maithel SK, Roa JC, Adsay V, Coimbra FJ, Jarnagin WR (2015) Gallbladder cancer: expert consensus statement. *HPB (Oxford)* 17:681–690. <https://doi.org/10.1111/hpb.12444>
- Lee SE, Kim KS, Kim WB, Kim IG, Nah YW, Ryu DH, Park JS, Yoon MH, Cho JY, Hong TH, Hwang DW, Choi DW (2014) Practical guidelines for the surgical treatment of gallbladder cancer. *J Korean Med Sci* 29:1333–1340. <https://doi.org/10.3346/jkms.2014.29.10.1333>
- National Comprehensive Cancer Network. The NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines™) Hepatobiliary Cancers. (Version 1.2021). 2021.
- Valle JW, Borbath I, Khan SA, Huguet F, Gruenberger T, Arnold D, Committee EG (2016) Biliary cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 27:v28–v37. <https://doi.org/10.1093/annonc/mdw324>
- Kim HS, Park JW, Kim H, Han Y, Kwon W, Kim SW, Hwang YJ, Kim SG, Kwon HJ, Vinueza E, Járufe N, Roa JC, Han IW, Heo JS, Choi SH, Choi DW, Ahn KS, Kang KJ, Lee W, Jeong CY, Hong SC, Troncso A, Losada H, Han SS, Park SJ, Yanagimoto H, Endo I, Kubota K, Wakai T, Ajiki T, Adsay NV, Jang JY (2018) Optimal surgical treatment in patients with T1b gallbladder cancer: An international multicenter study. *J Hepatobiliary Pancreat Sci* 25:533–543. <https://doi.org/10.1002/jhbp.593>
- Lee H, Kwon W, Han Y, Kim JR, Kim SW, Jang JY (2018) Optimal extent of surgery for early gallbladder cancer with regard to long-term survival: a meta-analysis. *J Hepatobiliary Pancreat Sci* 25:131–141. <https://doi.org/10.1002/jhbp.521>
- Lee SE, Jang JY, Kim SW, Han HS, Kim HJ, Yun SS, Cho BH, Yu HC, Lee WJ, Yoon DS, Choi DW, Choi SH, Hong SC, Lee

- SM, Kim HJ, Choi IS, Song IS, Park SJ, Jo S, Korean Pancreas Surgery C (2014) Surgical strategy for T1 gallbladder cancer: a nationwide multicenter survey in South Korea. *Ann Surg Oncol* 21:3654–3660. <https://doi.org/10.1245/s10434-014-3527-7>
11. Ouchi K, Mikuni J, Kakugawa Y, Organizing Committee TtA-CotJSoBS (2002) Laparoscopic cholecystectomy for gallbladder carcinoma: results of a Japanese survey of 498 patients. *J Hepatobiliary Pancreat Surg* 9:256–260. <https://doi.org/10.1007/s005340200028>
 12. Wakai T, Shirai Y, Yokoyama N, Nagakura S, Watanabe H, Hatakeyama K (2001) Early gallbladder carcinoma does not warrant radical resection. *Br J Surg* 88:675–678. <https://doi.org/10.1046/j.1365-2168.2001.01749.x>
 13. You DD, Lee HG, Paik KY, Heo JS, Choi SH, Choi DW (2008) What is an adequate extent of resection for T1 gallbladder cancers? *Ann Surg* 247:835–838. <https://doi.org/10.1097/SLA.0b013e3181675842>
 14. Kim BS, Ha HK, Lee IJ, Kim JH, Eun HW, Bae IY, Kim AY, Kim TK, Kim MH, Lee SK, Kang W (2002) Accuracy of CT in local staging of gallbladder carcinoma. *Acta Radiol* 43:71–76
 15. Kim SJ, Lee JM, Lee JY, Choi JY, Kim SH, Han JK, Choi BI (2008) Accuracy of preoperative T-staging of gallbladder carcinoma using MDCT. *AJR Am J Roentgenol* 190:74–80. <https://doi.org/10.2214/AJR.07.2348>
 16. Yoshimitsu K, Honda H, Shinozaki K, Aibe H, Kuroiwa T, Irie H, Chijiwa K, Asayama Y, Masuda K (2002) Helical CT of the local spread of carcinoma of the gallbladder: evaluation according to the TNM system in patients who underwent surgical resection. *AJR Am J Roentgenol* 179:423–428. <https://doi.org/10.2214/ajr.179.2.1790423>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.