Diagnostic accuracy of endoscopic ultrasound and intraductal ultrasonography for assessment of ampullary tumors: a meta-analysis

Xiaohua Ye\textsuperscript{a}, Lei Wang\textsuperscript{b} and Zhendong Jin\textsuperscript{b}

\textsuperscript{a}Department of Gastroenterology, Affiliated Jinhua Hospital, Zhejiang University School of Medicine, Jinhua, P.R. China; \textsuperscript{b}Department of Gastroenterology, Digestive Endoscopy Center, Changhai Hospital, Naval Medical University, Shanghai, P.R. China

**ABSTRACT**

**Background:** Accurate preoperative assessment of ampullary tumors (ATs) is critical for determining the appropriate treatment. The reported diagnostic accuracy of endoscopic ultrasound (EUS) and intraductal ultrasonography (IDUS) for detecting tumor depth (T-staging) and regional lymph node status (N-staging) varies across studies.

**Method:** An electronic search of the MEDLINE and Embase databases was conducted to identify studies that assessed the diagnostic accuracy of EUS and IDUS for ATs. Sensitivities and specificities of eligible studies were summarized using either fixed effects or random-effects model.

**Results:** Twenty-one studies were included in the final analysis. The pooled sensitivity and specificity of EUS were 0.89 and 0.87 for T1, 0.76 and 0.91 for T2, 0.81 and 0.94 for T3 and 0.72 and 0.98 for T4, respectively. For IDUS, estimates from five studies were 0.90 and 0.88 for T1, 0.73 and 0.91 for T2 and 0.79 and 0.97 for T3, respectively. For N-staging, 16 studies using EUS were included with sensitivity and specificity of 0.61 and 0.77, respectively. Moreover, estimates of IDUS for N-staging were 0.61 and 0.92, respectively.

**Conclusion:** Our results imply that EUS and IDUS have good diagnostic accuracy for T-staging of ATs. However, the accuracy of EUS or IDUS is less satisfactory for N-staging. More well-designed prospective studies are warranted to confirm our findings.

**Introduction**

Ampullary tumors (ATs) originate from the ampulla of Vater itself, distal to the bifurcation of the distal common bile duct and the pancreatic duct [1]. ATs have been increasingly diagnosed over the last decades, due to the wide use of endoscopic and radiological modalities for unrelated or other indications [2].

The removal of ATs is recommended in most cases because of its malignant attribute, especially when symptoms are present [1]. However, radical surgery carries high mortality rates ranging from 0 to 13% and high morbidity rates ranging from 25 to 63% [3]. In contrast, endoscopic ampullectomy appears to be a valid alternative to surgery for ATs. Despite the high rate of radical resections and low recurrence rate, the incidence of adverse events such as pancreatitis and hemorrhage should not be neglected [4]. A meta-analysis have showed that the overall rate of adverse events was up to 24.9% [5]. Therefore, an accurate preoperative assessment of ATs is crucial for triage of patients to endoscopic or surgical treatment.

The recent published European Society of Gastrointestinal Endoscopy (ESGE) guideline recommended endoscopic ultrasound (EUS) and intraductal ultrasonography (IDUS) for locoregional staging of ATs with low quality of evidence [6]. The main advantage of EUS is that the transducer can be placed close to the lesion without interference of fat, bowel gas or bone. IDUS provides real-time, high-quality cross-sectional images, and previous studies indicated that IDUS offers diagnostic yields that are equivalent to or slightly greater than those of EUS [7,8].

Various studies have evaluated the diagnostic accuracy of EUS or IDUS in endosonographic evaluation of T- and N-staging of ATs. However, results from these studies vary considerably. To our knowledge, a meta-analysis published in 2014 investigated the accuracy of EUS alone in ATs and found that EUS had a moderate strength of agreement with histopathology in determining T- and N-staging [9]. Since almost a decade has passed, new studies concerning the diagnostic accuracy of EUS or IDUS have been published, adding new information to the body of evidence. Thus, we conducted a systematic review and meta-analysis to update the current evidence.

**Methods**

**Search strategy**

We followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses and guidelines from Cochrane...
Diagnostic Test Accuracy Working Group to perform this meta-analysis [10]. Two investigators (Ye XH and Wang L) independently performed a computerized search of MEDLINE (from 1 January 1966 to 31 December 2021) and Embase (from 1 January 1974 to 31 December 2021) databases to identify potentially relevant articles. The search was carried out using the following keywords: endosonography (‘endoscopic ultrasound’, ‘EUS’, ‘intraductal ultrasonography’ and ‘IDUS’), ampullary (‘ampulla’ and ‘papilla’) and tumor (‘malignancy’, ‘neoplasm’, ‘cancer’ and ‘adenoma’). Manual searches of the bibliographies from these potential articles were also performed to identify additional studies.

**Study selection**

Two investigators (Ye XH and Wang L) independently reviewed potentially relevant articles for eligibility and inclusion. Studies were included if they met the following inclusion criteria: (1) Retrospective or prospective design published in manuscript form; (2) studies involving 10 or more patients using EUS or IDUS to evaluate ATs; (3) an appropriate reference standard was reported (endoscopic or surgical pathology); (4) reported absolute numbers of true-positive, false-negative, true-negative and false-positive observations for ATs, or if sufficient data could construct a 2 × 2 contingency table and (5) ATs were evaluated according to Tumor Node Metastasis (TNM) classification [11]. Case reports, editorials, review articles or clinical guidelines were excluded. Any disagreements were resolved by consensus.

For T-staging, T1 refers to lesion limited to the ampulla of Vater or sphincter of Oddi; T2 refers to invasion of the duodenal muscularis propria/duodenal wall; T3 refers to invasion of the pancreas and T4 refers to invasion of the peripancreatic soft tissue or adjacent organs or structures other than the pancreas. For N-staging, regional lymph nodes were defined as N1 if there were malignant regional lymph nodes on surgical pathology and N0 if no malignant regional lymph nodes were detected.

**Data extraction and quality assessment**

A custom-made standardized form was used for data extraction. For each eligible study, the following data were extracted: surname of first author, publication year, region of the study population, study design, sample size (the number of patients with ATs), details of the endosonographic type (EUS or IDUS, radial or linear) and reference standard (surgical or endoscopic pathology).

The methodological qualities of the studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [12]. This tool consists of four key categories: patient selection, index test, reference standard and flow and timing. Each category was assessed in terms of risk of bias, and the first three categories were also considered in terms of applicability.

**Data synthesis and statistical analysis**

The 2 × 2 tables (numbers of true-positive, false-negative, true-negative and false-positive) were constructed based on the data of the including studies. The pooled sensitivity and specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR) were calculated [13]. The summary estimates of sensitivities and specificities along with their corresponding 95% confidence intervals (CI) and prediction region were presented with a summary receiver-operating characteristic curve. Moreover, the area under the curve (AUC) was calculated [14]. Most clinical tests have an AUC value between 0.5 and 1.0, with a better diagnostic performance correlating with an AUC closer to 1.0 [15].

Heterogeneity was assessed using the Q-statistic and quantified using $I^2$. For the Q test, $p < .10$ was considered to imply statistical heterogeneity. $I^2$ is the proportion of total variation contributed by between-study variation. Deek’s test was used to evaluate publication bias [16]. All statistical analyses were carried out using Meta-Disc (version 1.4; Unit of Clinical Biostatistics, Ramony Cajal Hospital, Madrid, Spain) and STATA software (version 12.0; College Station, Texas, USA). A $p$ value less than .05 was considered statistically significant.

**Results**

**Study selection and characteristics**

Our initial search identified 1242 articles, and 1206 were subsequently excluded by examination of the titles or abstracts because they were reviews, experimental studies, meta-analyses, comments and other irrelevant articles. Of the remaining 36 articles, 15 articles were subsequently excluded from the meta-analysis: 5 did not provide sufficient data to construct a 2 × 2 table, 6 did not focus on the diagnostic study of interest, 1 was based on the same patient cohort, 1 included fewer than 10 patients and 2 reported miscellaneous neoplasms other than ATs. As a result, 21 studies [8,17–36] were identified in the final meta-analysis (Figure 1).

Baseline characteristics of the included studies are summarized in Table 1. The eligible studies were published between 1988 and 2019 and were performed in 12 regions. The sample size in each study ranged from 12 to 120 and included a total of 736 cases. Based on the QUADAS-2 tool, study bias and applicability outcomes were assessed, and the results are shown in Figure 2. Thirteen of the 21 studies were judged as high or unclear risk in one or more of the four key categories.

**Accuracy of EUS and IDUS to T-staging**

Eighteen studies evaluated the diagnostic performance of EUS on T-staging of ATs [18–25,27–36]. The pooled sensitivity and specificity of EUS were 0.89 (n [number of studies] = 15, 95%CI 0.85 – 0.92) and 0.87 (95%CI 0.83 – 0.91) for T1. For T2, EUS had a pooled sensitivity and specificity of 0.76.
1242 potentially relevant citations were initially retrieved

Articles were excluded based on titles and abstracts (n = 1206)

36 articles were eligible for further evaluation

Insufficient data to construct a 2×2 table (n = 5)
Did not focus on the diagnostic study of interest (n = 6)
Based on the same patient cohort (n = 1)
Included fewer than 10 patients (n = 1)
Reported miscellaneous neoplasms other than ATs (n = 2)

21 articles finally met the inclusion criteria

Figure 1. Flow chart of the study selection.

Table 1. Baseline characteristics of included studies in meta-analysis.

| Author (Year) | Region          | Study design | Sample size | Diagnostic method | Reference standard   |
|---------------|-----------------|--------------|-------------|-------------------|----------------------|
| Yasuda et al. 1988 [17] | Japan           | Retrospective | 12          | Radial            | Surgical pathology   |
| Mukai et al. 1992 [18]  | Japan           | Retrospective | 23          | Radial            | Surgical pathology   |
| Rattner et al. 1996 [19] | US              | Retrospective | 12          | Radial or linear  | Surgical pathology   |
| Tio et al. 1996 [20]    | Netherlands     | Retrospective | 32          | Radial            | Surgical pathology   |
| Zhang et al. 1996 [21]  | China           | Retrospective | 22          | Radial            | Surgical pathology   |
| Itoh et al. 1997 [22]   | Japan           | Retrospective | 32          | Radial and IDUS   | Surgical pathology   |
| Sauvanet et al. 1997 [23]| France         | Retrospective | 26          | Radial            | Surgical pathology   |
| Cannon et al. 1999 [24] | US              | Retrospective | 50          | Radial or linear  | Surgical pathology   |
| Kubo et al. 1999 [25]   | Japan           | Retrospective | 35          | Radial            | Surgical pathology   |
| Menzel et al. 1999 [26] | Germany         | Prospective   | 15          | IDUS              | Surgical pathology   |
| Skordilis et al. 2002 [27]| Greece         | Prospective   | 17          | Radial            | Surgical pathology   |
| Ito et al. 2007 [28]    | Japan           | Prospective   | 40          | Radial and IDUS   | Surgical or endoscopic pathology |
| Artifon et al. 2009 [29] | Brazil and US  | Prospective   | 27          | Linear            | Surgical pathology   |
| Chen et al. 2009 [30]   | Taiwan          | Retrospective | 33          | Radial            | Surgical pathology   |
| Manta et al. 2010 [31]  | Italy           | Retrospective | 24          | Linear            | Surgical or endoscopic pathology |
| Heinzow et al. 2011 [8] | Germany         | Retrospective | 28          | IDUS              | Surgical pathology   |
| Wee et al. 2012 [32]    | India           | Retrospective | 26          | Radial or linear  | Surgical pathology   |
| Okano et al. 2013 [33]  | Japan           | Retrospective | 48          | Radial and IDUS   | Surgical or endoscopic pathology |
| Riddit et al. 2014 [34] | US              | Retrospective | 100         | Radial or linear  | Surgical pathology   |
| Membrillo-Romero et al. 2019 [35] | Mexico   | Retrospective | 14          | Radial or linear  | Surgical pathology   |
| Peng et al. 2019 [36]   | China           | Retrospective | 120         | Radial or linear  | Surgical or endoscopic pathology |

IDUS: intraductal ultrasonography.

(n = 17, 95%CI 0.69 – 0.82) and 0.91 (95%CI 0.88 – 0.94). Estimates of sensitivity and specificity were 0.81 (n = 14, 95%CI 0.73 – 0.88) and 0.94 (95%CI 0.91 – 0.96) in determining T3 stage tumor. To diagnose T4 stage tumor, EUS had a pooled sensitivity of 0.72 (n = 6, 95%CI 0.47 – 0.90) and specificity of 0.98 (95%CI 0.96 – 0.99), respectively. Forest plots of the pooled sensitivity and specificity for EUS are shown in Figure 3. The AUC of EUS for T-staging ranged from 0.89 to 0.95 (Figure 4 and Table 2). The calculated I² value is over 50% in sensitivity for T1 (72.4%) and specificity for T1 (66.7%), T2 (68.6%) and T3 (60.8%) (Figure 3). The PLR, NLR and DOR values of EUS to diagnose various T stages of ATs are shown in Table 2.

With regard to IDUS, five studies were included for evaluation of T-staging [8,22,26,28,33]. The summary results of sensitivity and specificity were 0.90 (n = 5, 95%CI 0.82 – 0.95) and 0.88 (95%CI 0.78 – 0.95) for T1, 0.73 (n = 5, 95% CI 0.54 – 0.88) and 0.91 (95%CI 0.85 – 0.95) for T2 and
0.79 (n = 2, 95%CI 0.49 – 0.95) and 0.97 (95%CI 0.84 – 1.00) for T3, respectively. The Forrest plots in Figure 5 show the sensitivity and specificity of IDUS in diagnosing various T stages. The AUC curves of IDUS for T-staging are shown in Figure 6. A test of heterogeneity for all the pooled estimates of IDUS had a p value > .10 expect for T2 in specificity. PLR, NLR and DOR of IDUS for various T stages are shown in Table 3.

Furthermore, two studies included T3 and T4 stage tumors in the same group and therefore additional analyses for T3-4 were performed [28,33]. The summary results of sensitivity and specificity were 0.80 (n = 2, 95%CI: 0.64 – 0.91) and 0.91 (95%CI: 0.82 – 0.96) for EUS and 0.80 (n = 2, 95%CI: 0.44 – 0.97) and 0.93 (95%CI: 0.84 – 0.97) for IDUS, respectively. The heterogeneity was significant for EUS and IDUS in specificity (p < .05). The AUC, PLR, NLR and DOR values of EUS and IDUS to diagnose T3-4 stage of ATs are shown in Tables 2 and 3.

Accuracy of EUS and IDUS to N-staging

To differentiate N-staging, EUS had a pool sensitivity of 0.61 (n = 16, 95%CI: 0.53 – 0.69) and a specificity of 0.77 (95%CI: 0.71 – 0.82), respectively [17–25,27,29–32,35,36]. For IDUS, the summarized sensitivity and specificity were 0.61 (n = 3, 95%CI: 0.39 – 0.80) and 0.92 (95%CI: 0.81 – 0.98) (Figure 7). The pooled AUC values of EUS and IDUS for N-staging were 0.74 and 0.87, respectively (Figure 8). p Values for the Chi-squared test were >.10 for all pooled estimates except for EUS in specificity. PLR, NLR and DOR of EUS and IDUS for various N stages are shown in Tables 2 and 3, respectively.

Accuracy of EUS and IDUS to intraductal extension

The overall pooled sensitivity and specificity of EUS in diagnosing intraductal (pancreatic or biliary duct) extension were 0.79 (n = 3, 95%CI: 0.64 – 0.91) and 0.88 (95%CI: 0.83 – 0.92)
The AUC value of EUS for intraductal extension was 0.92. The values of PLR, NLR and DOR were 5.39, 0.3 and 18.33, respectively. Only one study separately investigated the efficacy of IDUS for pancreatic and biliary duct extension [33]. For pancreatic duct extension, the sensitivity and specificity were 0.14 (95%CI: 0.04 – 0.58) and 0.99 (95%CI: 0.91 – 1.00), respectively. The estimates for biliary duct extension were 0.56 (95%CI: 0.25 – 0.82) and 0.95 (95%CI: 0.82 – 0.99), respectively. The AUC values of IDUS were 0.57 for pancreatic duct extension and 0.75 for biliary duct extension.

**Figure 3.** Forrest plots of sensitivity and specificity of EUS for T-staging of ATs. EUS: endoscopic ultrasound; ATs: ampullary tumors.
The PLR, NLR and DOR values for pancreatic duct extension were 15.75, 0.82 and 19.15, whereas the results were 10.83, 0.47 and 23.13 for biliary duct extension.

Assessment of heterogeneity

Heterogeneity was mainly found in sensitivity (T1) and specificity (T1, T2, T3 and N-staging) of the summarized EUS results. Subgroup analyses were performed based on publication year (before 2000 vs. after 2000), area (eastern countries vs. western countries), EUS technique (radial EUS only vs. radial or linear EUS) and study design (retrospective vs. prospective). For T1, the heterogeneity in sensitivity was attributed to EUS technique and study design. The heterogeneity in specificity (T1, T2, T3 and N-staging) was eliminated or reduced to varying degrees when subgrouping based on publication year, area, EUS technique and study design, suggesting that these factors might be significant contributors (Supplementary Table 1).

### Table 2. Diagnostic performance of EUS in T/N staging for ATs.

|       | T1                  | T2                  | T3  | T4                  | T3–4 | N  |
|-------|---------------------|---------------------|-----|---------------------|------|----|
| No. of studies (patients) | 15 (623)            | 17 (655)            | 14 (543) | 6 (345)            | 2 (88) | 16 (390) |
| Sensitivity (95%CI)      | 0.89 (0.85–0.92)    | 0.76 (0.69–0.82)    | 0.81 (0.73–0.88) | 0.72 (0.47–0.90) | 0.80 (0.44–0.97) | 0.61 (0.53–0.69) |
| Specificity (95%CI)      | 0.87 (0.83–0.91)    | 0.91 (0.88–0.94)    | 0.94 (0.91–0.96) | 0.98 (0.96–0.99) | 0.91 (0.82–0.96) | 0.77 (0.71–0.82) |
| PLR (95%CI)              | 5.76 (3.48–9.52)    | 4.88 (3.25–7.33)    | 7.08 (4.88–10.27) | 19.25 (8.26–44.86) | 8.87 (0.91–86.58) | 2.44 (1.50–3.96) |
| NLR (95%CI)              | 0.26 (0.17–0.39)    | 0.39 (0.31–0.50)    | 0.26 (0.19–0.37) | 0.46 (0.26–0.82) | 0.34 (0.10–1.13) | 0.61 (0.50–0.75) |
| DOR (95%CI)              | 28.31 (15.78–50.81) | 20.99 (12.24–35.98) | 39.20 (20.43–75.21) | 66.51 (16.88–261.96) | 36.27 (5.37–245.07) | 4.67 (2.63–8.28) |
| AUC (SE)                 | 0.91 (0.02)         | 0.89 (0.02)         | 0.93 (0.02) | 0.95 (0.03) | 0.91 (0.04) | 0.74 (0.04) |

EUS: endoscopic ultrasound; ATs: ampullary tumors; CI: confidence interval; PLR: positive likelihood ratio; NLR: negative likelihood ratio; AUC: area under curve; SE: standard error; DOR: diagnostic odds ratio.

**Assessment of heterogeneity**

Heterogeneity was mainly found in sensitivity (T1) and specificity (T1, T2, T3 and N-staging) of the summarized EUS results. Subgroup analyses were performed based on publication year (before 2000 vs. after 2000), area (eastern countries vs. western countries), EUS technique (radial EUS only vs. radial or linear EUS) and study design (retrospective vs. prospective). For T1, the heterogeneity in sensitivity was attributed to EUS technique and study design. The heterogeneity in specificity (T1, T2, T3 and N-staging) was eliminated or reduced to varying degrees when subgrouping based on publication year, area, EUS technique and study design, suggesting that these factors might be significant contributors (Supplementary Table 1).
Publication bias

The publication bias was assessed using Deek’s funnel plot. If exists, publication bias results in a higher proportion of smaller studies with bigger effect sizes compared to larger ones. In the funnel plot, the vertical axis represents the inverse of the square root of the effective sample size, while the horizontal axis represents the DOR. With the exception of EUS for T4 (p < .01), all the other Deek’s funnel plots were symmetrical with respect to the regression line, and
asymmetry tests revealed no evidence of publication bias. The funnel plots to investigate the effect of publication bias for EUS and IDUS estimating T- and N-staging of ATs are shown in Figure 9.

**Discussion**

In this study, we conducted a robust systematic review and an appropriately performed meta-analysis. Our main findings were that both EUS and IDUS had acceptable sensitivities (0.72 – 0.89 for EUS; 0.73 – 0.90 for IDUS) and specificities (0.87 – 0.98 for EUS; 0.88 – 0.97 for IDUS) in diagnosing T-staging of ATs, whereas the accuracy of EUS or IDUS is less satisfactory for N-staging. The AUC values of EUS (0.89 – 0.95) and IDUS (0.87 – 0.95) for T-staging were very close to 1, indicating that both EUS and IDUS are excellent T-staging tests for ATs. With regard to N-staging, the summarized results suggest that either EUS (sensitivity, 0.61; specificity, 0.77; AUC, 0.74) or IDUS (sensitivity, 0.61; specificity, 0.92; AUC, 0.87) is suboptimal. In consideration of intraductal extension as a predictor for incomplete endoscopic resection and recurrence [34,37], data were also collected. According to our synthesized results, EUS has been shown to be useful in evaluating intraductal extension of ATs (sensitivity, 0.79; specificity, 0.88; AUC, 0.92).
To the best of our knowledge, this is the first meta-analysis that quantitatively summarizes all the available evidence of both EUS and IDUS in the locoregional staging of ATs. A previous systematic review was published in 2014, which included 14 studies with respect to the diagnostic accuracy of EUS alone on ATs [9]. In our study, we included 21 studies and strengthened the body of evidence. Several included studies reported results for T-staging as ‘T3-4’, which precluded data extraction for T-staging. However, the prior systematic review considered ‘T3-4’ as T3 and T4 separately.

Figure 9. Funnel plots assessing bias for T- and N-staging of EUS and IDUS for ATs. EUS: endoscopic ultrasound; IDUS: intraductal ultrasonography; ATs: ampullary tumors.
The main strength of our meta-analysis is that we performed additional analyses of studies that reported T-staging as ‘T3-4’ and found that the performance of EUS and IDUS to diagnose T3-4 tumors was comparable with other results, hence improving the estimate accuracy.

The incidence of ATs has increased in clinical practice due to the development of routine screening endoscopic procedures and imaging modalities [2]. Consequently, accurate locoregional assessment is of great importance for selecting optimal treatment modality. The ESGE guideline published in 2021 has recommended EUS and IDUS for locoregional staging of ATs; however, the quality of evidence was low [6]. In addition, ESGE guideline also stipulated that other imaging modalities such as abdominal magnetic resonance cholangiopancreatography (MRCP) for staging of ATs. Generally, various imaging technologies including CT scan, MRCP and transabdominal ultrasound are traditionally used combined with EUS for preoperative staging of ATs in clinical practice. Indeed, several studies suggest that EUS provides significantly higher performance specifically for T-staging compared with CT and transabdominal ultrasound, whereas comparable or slightly higher accuracy compared with MRCP with no statistical significance [17,18,24,25,38]. Lymph node metastasis is a well-established prognostic predictor for ATs [36]. Although the accuracy is not as reliable as that of T-staging, EUS and IDUS can still help clinicians stratify the risk of patients with lymph node metastatic disease and represent a clue in selecting patients for optimal treatments. In addition, studies have demonstrated that the performance of EUS was not statistically different as compared to MRCP and CT [22,30,31,33]. Recently, several reports have described EUS-guided fine-needle aspiration (FNA) for ATs and it might be another diagnostic option with a sensitivity of 86.0% and a specificity of 100% [39,40].

IDUS uses higher frequency ultrasound probe (20–30 MHz) and thus produces higher-resolution images than EUS [41]. IDUS provides superior differentiation between the sphincter of Oddi and the duodenal wall because less tissue is compressed during scanning. Previous studies demonstrate that IDUS has slightly higher diagnostic yields than or comparable to those of EUS [7,8] and is helpful in selecting appropriate patients for indication of endoscopic ampullectomy [28]. Moreover, as the guideline recommended, routine use of IDUS should weigh against training, costs and risk of pancreatitis [6]. However, the number of included cases with ATs undergoing IDUS was limited, and larger series with longer follow up are imperative for exploring its clinical significance.

Our data carry clinical implications. As shown in Tables 2 and 3, DOR, PLR and NLR were calculated. DOR refers to the odds of having a positive test in patients with a true historical stage of the disease when compared with patients who do not have the disease. For instance, the odds of having the correct histological T1 stage of ATs is 28.31 times. It enables physicians determining treatment strategy with confidence. PLR defines as a measure of how well the test correctly identifies a disease, whereas NLR is a measure of how well the test correctly excludes a disease [42]. Comparatively speaking, likelihood ratios including PLR and NLR are supposed to be more clinically practical. The results indicate that both EUS and IDUS perform well in excluding as well as diagnosing the correct T stage of ATs.

Some limitations of our study merit consideration. First, most of our studies were retrospective in design, thereby may overestimating the diagnostic precision. Also, some of the studies did not specifically differentiate benign ampullary adenomas from ampullary cancers. Third, the definition of lymph node metastasis varied across studies, thereby leading to selection bias. Fourth, for IDUS results, there was only a small number of studies and cases included to draw a robust conclusion. Therefore, the diagnostic performance of IDUS for ATs might be less reliable. The resolution of this issue will require more data and additional studies. Last, it is well known that intraobserver variability exists in EUS and IDUS interpretation and thus may affect the accuracy of our analyses.

In summary, EUS and IDUS are highly accurate techniques in T-staging and intraductal extension for ATs. However, EUS and IDUS are less satisfactory for N-staging due to their modest sensitivities and specificities. The results of the test must be interpreted with caution in specific clinical contexts. More prospective, well-designed studies are demanded.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This work was supported by Natural Science Foundation of Zhejiang Province (LQ19H030003) and Key Project of Jinhua Science and Technology Bureau (2018A32022).

References
[1] Poley JW, Campos S. Methods and outcome of the endoscopic treatment of ampullary tumors. Ther Adv Gastrointest Endosc. 2020;13:2631774519899786.
[2] Saha SK, Zhu AX, Fuchs CS, et al. Forty-year trends in cholangiocarcinoma incidence in the U.S.: intrahepatic disease on the rise. Oncologist. 2016;21(5):594–599.
[3] Klair JS, Irani S, Kozarek R. Best techniques for endoscopic ampullectomy. Curr Opin Gastroenterol. 2020;36(5):385–392.
[4] Meunier C, Lisotti A, Gupta V, et al. Oral anticoagulants but not antiplatelet agents increase the risk of delayed bleeding after endoscopic papillotomy: a large study in a tertiary referral center. Surg Endosc. 2022. DOI:10.1007/s00464-022-09138-5
[5] Spadaccini M, Fugazza A, Frazzoni L, et al. Endoscopic papillectomy for neoplastic ampullary lesions: a systematic review with pooled analysis. United Eur Gastroenterol J. 2020;8(1):44–51.
[6] Vanbierdeghe T, Strikker M, Arvanitakis M, et al. Endoscopic management of ampullary tumors: European society of gastrointestinal endoscopy (ESGE) guideline. Endoscopy. 2021;53(4):429–448.
[7] Mengel J, Pretorius M, Dietl EH, et al. Tumors of the papilla of Vater—adequate diagnostic impact of endoscopic forceps biopsies taken prior to and following sphincterotomy. Ann Oncol. 1999;10(10):1227–1231.
[8] Heinzow HS, Lenz P, Lallier S, et al. Ampulla of Vater tumors: impact of intraductal ultrasound and transpapillary endoscopic biopsies on diagnostic accuracy and therapy. Acta Gastroenterol Belg. 2011;74(4):509–515.
Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–536.

Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982–990.

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20(19):2865–2884.

Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg. 2005;79(1):16–20.

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58(9):882–893.

Yasuda K, Mukai H, Cho E, et al. The use of endoscopic ultrasonography in the diagnosis and staging of carcinoma of the papilla of Vater. Endoscopy. 1988;20 Suppl 1(Suppl 1):218–222.

Mukai H, Nakajima M, Yasuda K, et al. Evaluation of endoscopic ultrasonography in the pre-operative staging of carcinoma of the ampulla of Vater and common bile duct. Gastrointest Endosc. 1992;38(6):676–683.

Rattner DW, Fernandez-del Castillo C, Brugge WR, et al. Defining the criteria for local resection of ampullary neoplasms. Arch Surg. 1996;131(4):366–371.

Tio TL, Sie LH, Kallimanis G, et al. Staging of ampullary and pancreatic carcinoma: comparison between endosonography and surgery. Gastrointest Endosc. 1996;44(6):706–713.

Zhang QL, Nian WD, Zhang LP, et al. Endoscopic ultrasonography assessment for ampullary and bile duct malignancy. Diagn Ther Endosc. 1996;3(1):35–40.

Itoh A, Goto H, Naitoh Y, et al. Intraductal ultrasonography in diagnosing tumor extension of cancer of the papilla of Vater. Gastrointest Endosc. 1997;45(3):251–260.

Sauvanet A, Chapuis O, Hammel P, et al. Are endoscopic procedures able to predict the benignity of ampullary tumors? Am J Surg. 1997;174(3):355–358.

Cannon ME, Carpenter SL, Elta GH, et al. EUS compared with CT, magnetic resonance imaging, and angiography and the influence of biliary stenting on staging accuracy of ampullary neoplasms. Gastrointest Endosc. 1999;50(1):27–33.

Kubo H, Chijiwa Y, Akahoshi K, et al. Pre-operative staging of ampullary tumors by endoscopic ultrasound. Br J Radiol. 1999;72(857):443–447.

Menzel J, Hoepffner N, Sulkowski U, et al. Polypoid tumors of the major duodenal papilla: preoperative staging with intraducal US, EUS, and CT—a prospective, histopathologically controlled study. Gastrointest Endosc. 1999;49(3 Pt 1):349–357.

Sكورديس P, Mouzas IA, Dimoulias PD, et al. Is endosonography an effective method for detection and local staging of the ampullary carcinoma? A prospective study. BMC Surg. 2002;2(1):1.

Ito K, Fujita N, Noda Y, et al. Preoperative evaluation of ampullary neoplasm with EUS and transpapillary intraductal US: a prospective and histopathologically controlled study. Gastrointest Endosc. 2007;66(4):740–747.

Artifon EL, Couto D Jr, Sakai P, et al. Prospective evaluation of EUS versus CT scan for staging of ampullary cancer. Gastrointest Endosc. 2009;70(2):290–296.

Chen CH, Yang CC, Yeh YH, et al. Reappraisal of endosonography of ampullary tumors: correlation with transabdominal sonography, CT, and MRI. J Clin Ultrasound. 2009;37(1):18–25.

Manta R, Conigliaro R, Castellani D, et al. Linear endoscopic ultrasonography vs magnetic resonance imaging in ampullary tumors. World J Gastroenterol. 2010;16(44):5592–5597.

Wee E, Lakhtakia S, Gupta R, et al. The diagnostic accuracy and strength of agreement between endoscopic ultrasound and histopathology in the staging of ampullary tumors. Indian J Gastroenterol. 2012;31(6):324–332.

Okano N, Igarashi Y, Hara S, et al. Endoscopic preoperative evaluation for tumors of the ampulla of Vater using endoscopic ultrasonography and intraductal ultrasonography. Clin Endosc. 2014;47(2):174–177.

Ridtitid W, Schmidt SE, Al-Haddad MA, et al. Performance characteristics of EUS for locoregional evaluation of ampullary lesions. Gastrointest Endosc. 2015;81(2):380–388.

Membrillo-Romero A, Fuente-Lira M, Carvallo-Guevara P, et al. Ampullary tumors: presurgical staging by endoscopic ultrasonography. Diagnostic test study. Cir Cir. 2019;87(3):278–284.

Peng CY, Lv Y, Shen SS, et al. The impact of endoscopic ultrasound in preoperative evaluation for ampullary adenomas. J Dig Dis. 2019;20(5):248–255.

Kawashima H, Ohno E, Ishikawa T, et al. Endoscopic papillectomy for ampullary adenoma and early adenocarcinoma: analysis of factors related to treatment outcome and long-term prognosis. Dig Endosc. 2021;33(5):858–869.

Chen CH, Tseng LJ, Yang CC, et al. Preoperative evaluation of periampullary tumors by endoscopic sonography, transabdominal sonography, and computed tomography. J Clin Ultrasound. 2001;29(6):313–321.

Defrain C, Chang CY, Srikureja W, et al. Cytologic features and diagnostic pitfalls of primary ampullary tumors by endoscopic ultrasound-guided fine needle aspiration biopsy. Cancer. 2005;105(5):289–297.

Ogura T, Hara K, Hijioka S, et al. Can endoscopic ultrasound-guided fine needle aspiration offer clinical benefit for tumors of the ampulla of Vater? - an initial study. Endosc Ultrasound. 2012;1(2):84–89.

Sun B, Hu B. The role of intraductal ultrasonography in pancreaticobiliary diseases. Endosc Ultrasound. 2016;5(5):291–299.

Puli SR, Bechtold ML, Reddy JB, et al. How good is endoscopic ultrasound in differentiating various T stages of rectal cancer? Meta-analysis and systematic review. Ann Surg Oncol. 2009;16(2):254–265.