Meta-analysis of the diagnostic performance of serum carbohydrate antigen 19-9 for the detection of gallbladder cancer

Xiaolei Zhou

Abstract

Introduction: Carbohydrate antigen 19-9 (CA19-9) is a well-studied tumor marker, yet its diagnostic value for gallbladder cancer remains unclear. The present meta-analysis was conducted to validate the role of serum CA19-9 for the detection of gallbladder cancer.

Methods: A systematic search of digital databases was conducted, complemented by additional hand-searching. Studies that reported serum CA19-9 for the differentiation of gallbladder cancer cases from non-gallbladder cancer controls were considered eligible.

Results: A total of 27 studies involving 4300 subjects were included. The pooled sensitivity, specificity, and area under the curve in diagnosing gallbladder cancer were 0.70 (95% confidence interval (CI): 0.63–0.76), 0.92 (95% CI: 0.88–0.94), and 0.89 (95% CI: 0.86–0.92), respectively. The pooled positive likelihood rate, negative likelihood rate, and diagnostic odds rate were 8.30 (95% CI: 5.84–11.69), 0.33 (95% CI: 0.27–0.41), and 25.13 (95% CI: 5.83–39.89), respectively. Meta-regression analysis revealed that there was significantly lower sensitivity (0.69, 95% CI: 0.61–0.77) and specificity (0.91, 95% CI: 0.87–0.95) when CA19-9 was used for the differentiation of gallbladder cancer cases from benign biliary diseases. A better specificity of 0.93 (95% CI: 0.90–0.96) was reached in the setting of a sample size ≥100.

Conclusions: Serum CA19-9 can be a potential candidate marker for the detection of gallbladder cancer, which maintains moderate sensitivity and good specificity. Attention should be paid to the control type and sample size, which may affect its diagnostic accuracy. Also, results should be interpreted with caution due to significant heterogeneity primarily caused by different thresholds between included studies.

Keywords

CA19-9 antigen, diagnosis, gallbladder cancer, meta-analysis, tumor marker

Date received: 3 November 2021; revised: 6 December 2021; accepted 7 December 2021

Introduction

Gallbladder cancer (GBC) is one of the most prevalent malignancies of the biliary system. It is generally considered rare but highly lethal. There were 84,695 fatal cases and 115,949 new diagnoses associated with GBC worldwide according to the Globocan 2020 Database. A patient with GBC has an abysmal prognosis as GBC is usually not diagnosed until it reaches an advanced stage where therapeutic strategies remain largely ineffective. The only potentially curative treatment for GBC remains surgical resection with clear margins; however, the lack of non-invasive approaches able to accurately detect the silent neoplasm at early stages might, in part, result in a
low curative resection rate. Moreover, it remains unclear whether current evidence is rigorous enough to support the role of adjuvant systematic treatment. An early diagnosis is therefore of great importance to develop optimal management options for patients with GBC and improve the outcomes.

A combination of clinical features, imaging techniques, and assessment of non-specific tumor markers is commonly employed for the detection of malignant tumors. Analysis of tumor markers is considered to be convenient, rapid, and economic compared to imaging examination. Despite various tumor markers having been evaluated, controversies abound regarding the validation of these existing markers for the diagnosis of gallbladder malignancy. Carbohydrate antigen 19-9 (CA19-9), also known as sialyl Lewis-a, is a well-studied serological tumor marker that was widely utilized primarily in the work-up of pancreatic carcinoma (PC) victims. Notably, CA19-9 has been shown to be a promising marker for GBC. Evidence has shown that patients with GBC had elevated CA19-9 levels in serum compared with those with benign gallbladder diseases and healthy individuals. Multiple studies conducted over a number of years have attempted to explain the proper use and interpretation of CA19-9 for the detection of GBC but yielded inconsistent results. Given that the diagnostic role of CA19-9 remains obscure, additional information is emboldened to establish a diagnostic profile of CA19-9 in this condition. Hence the present meta-analysis was undertaken to systematically collect findings to validate the diagnostic potential for the detection of GBC, thereby providing useful information for clinical practice.

Materials and methods

This meta-analysis followed the guidelines of 2020 preferred reporting items for systematic reviews and meta-analyses (PRISMA) statements. Research ethics and patient consent are not applicable for this meta-analysis.

Literature search

Computerized searching of mainstream databases including PubMed, Embase, and Web of Science, was carried out to identify potentially relevant studies up to July 2021. China databases, including the China National Knowledge Infrastructure (CNKI) and the Wanfan Database, were also queried with search patterns. The following keywords with their abbreviations and synonyms were employed in multiple pairwise combinations: “gallbladder”, “gallbladder cancer”, “biliary tract cancer”, “gastrointestinal carcinoma”, “malignant gallbladder lesions”, “cancer antigens”, “carbohydrate antigen”, “carbohydrate antigen 19-9”, “diagnosis”, “diagnostic marker”, “tumor marker”, and “diagnostic role”. A manual search of the references of potential articles was also performed. Following prior searches, the investigator removed duplicated entries with the help of Endnote software, and filtered the results according to titles and abstracts to select potential publications for further evaluation. A full-text examination as per pre-defined eligibility criteria was subsequently carried out to assess whether selected studies were eligible for the analysis.

Inclusion and exclusion criteria

Articles that met the following criteria were considered eligible: (a) participants: patients with confirmed GBC; (b) intervention: detection of CA19-9 in serum before treatment; (c) comparisons: non-GBC controls including benign biliary diseases and/or healthy individuals; (d) outcomes: reporting sufficient diagnostic data of CA19-9 so that true positive (TP), false positive (FP), true negative (TN) and false negative (FN) could be extracted or calculated to construct 2×2 contingency tables; and (e) study design: full-length articles written in English or Chinese. Exclusion criteria included: (a) incomplete data for the 2×2 table construction; and (b) less than 10 GBC cases. Except for original articles, other publications like reviews, conference abstracts, editorials, or book sections were rejected.

Data extraction and quality appraisal

The following data were extracted from each eligible study using a pre-established spreadsheet: the first author’s name, year of publication, time of admission, location, study design, number of enrolled subjects, patient characteristics (mean age, gender), control type, methods used for detecting CA19-9, summary measures of diagnostic performance, and cut-off values used to define positives of CA19-9. Regarding methodological quality, each study was assessed using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2) under Review Manager Version 5.3. This tool consists of two main categories: the risk of bias and applicability. The former contains four domains including patient selection, index test, reference standard, and flow and timing; for the latter, only the first three domains were evaluated. Each study was judged as having a high, unclear, or low risk of bias in terms of corresponding domains.

Statistical analysis

The elements for 2×2 table construction (TP, FP, FN, TN) extracted from original studies were applied as primary data to calculate the pooled sensitivity (SEN), specificity (SPE), positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR). PLR was the ratio of TPs and FPs. A higher PLR accompanied a better diagnostic performance. NLR was the ratio of TNs and FNs, with a lower NLR suggesting a higher diagnostic value. DOR was the ratio of PLR and NLR, used for a
general estimation of the efficacy of CA19-9 for the diagnosis of GBC. Summary receiver operating characteristics (SROC) curves were derived from observed data, with the area under the curve (AUC) of SROC calculated for a global evaluation of diagnostic performance. Data were analyzed using Stata Version 15.1 and a random-effects model was adopted for data pooling owing to expected heterogeneity. Inter-study heterogeneity was estimated by the Q statistics under the chi-square value test and the inconsistency index ($I^2$) test. An $I^2 > 50\%$ with a p-value $< 0.05$ suggested significant heterogeneity. The threshold effect could greatly contribute to heterogeneity for diagnostic meta-analyses. The existence of the threshold effect was therefore explored via the Spearman correlation coefficients between the logit of sensitivity and the logit of (1-SPE), and the proof of existence was provided by a strong positive correlation ($>0.6$) between sensitivity and (1-SPE) with a p-value $<0.05$. An additional sensitivity analysis was designed prospectively. The majority of studies were single-centered, while six studies adopted data from at least two centers for analysis. Table 1 shows detailed information on the main characteristics.

Results

Study identification

Database searches resulted in the initial identification of 898 records. After removal of duplicates, 720 articles underwent screening according to titles and abstracts. Subsequently, 148 were selected and considered possible contenders for eligibility. Two additional records were found through hand-searches and were assessed for inclusion. Following full-text evaluation, 27 studies were eventually included in the meta-analysis. A flow chart for study identification is shown in Figure S1.

Study characteristics

There were 20 studies written in Chinese and 7 in English. All were conducted in Asian countries, with 1 in Japan, 1 in India, and the rest in China. The overall number of enrolled subjects was 4300, consisting of 1516 patients with GBC and 2784 non-GBC controls (286 healthy individuals and 2498 patients with related benign diseases). The mean/median age of enrolled GBC cases ranged from 45 to 72. Among the included studies, the time of enrollment spanned 25 years (from 1994 to 2019). Regarding study design, only the study of Bind et al. was designed prospectively. The majority of studies were single-centered, while six studies adopted data from at least two centers for analysis. Table 1 shows detailed information on the main characteristics.

Quality evaluation

A high risk of bias was obviously seen in the domain of patient selection as most of the included papers were designed as prospective case-control studies, which enrolled subjects with confirmed GBC instead of suspected patients. Studies were rated as having a high or unclear risk of bias pertaining to an “index test” due to failure to mention pre-specified thresholds for diagnosis, or due to inexact reporting. Overall, the risk of bias was low in the other two domains. Also, there were no serious concerns regarding their applicability. A summary of the methodological quality of the included studies is provided in Figure S2.

Diagnostic performance

A random-effects model was adopted for data analysis owing to substantial heterogeneity (all $I^2 >50\%$, $p = 0.00$). Combined sensitivity and SPE of CA19-9 for the detection of GBC was 0.70 (95% confidence interval (CI): 0.63–0.76, Figure 1(a)), 0.92 (95% CI: 0.88–0.94, Figure 1(b)), respectively. The pooled PLR was 8.30 (95% CI: 5.84–11.69, $I^2$: 81.8%), and the pooled NLR was 0.33 (95% CI: 0.27–0.41, $I^2$: 89.7%). Pooled DOR of CA19-9 was 25.13 (95% CI: 15.83–39.89, $I^2$: 98.4%). An AUC of 0.89 (95% CI: 0.86–0.92) was obtained from SROC as shown in Figure 2, indicating that CA19-9 exhibited a moderate diagnostic accuracy in differentiating GBC from non-GBC, based on the observation that the AUC was in the range of 0.7 and 0.9. Also, the detection of publication bias was absent from the analysis, based upon a p-value of 0.56 under Deeks’ test (Figure S3).

A notable threshold effect was found in the analysis, implying that high heterogeneity was far more likely to result from the threshold effect (correlation: 0.63, $p = 0.02$). The ensuing sensitivity analysis for pooled results demonstrated that the studies of Bind et al., Wang et al., and Zhu et al. were the dominant studies in weight according to influence analysis (Cook’s distance $>1.00$; Figure S4(c)). Outlier detection identified that the studies of Bind et al. and Wang et al. might be the important contributors to considerable heterogeneity (Figure S4(d)). However, significant heterogeneity remained after excluding these studies one by one or in combination. Results are summarized in Table 2. Subgroup and meta-regression analyses were additionally performed, and summarized outcomes are displayed in Table 3. Control type and sample size were shown to significantly affect the SPE ($p < 0.05$), while only the control type influenced the sensitivity ($p < 0.05$). The diagnostic value of serum CA19-9 diminished when it was used to differentiate GBC from benign biliary diseases, and its SPE was improved when it was used in the setting of sample size $\geq 100$. 
Discussion

GBC is a form of hepatobiliary malignancy with an overall poor prognosis. Recognition of GBC in the early stage enables more effective treatment and improves prognosis, yet early diagnosis remains a challenging task due to the lack of specific symptoms and reliable markers. As a cancer-associated antigen, CA19-9 is the most intensively

Table 1. Summary of the main characteristics among included papers.

| 1st author | Year | Ref. standard | Case/Control | Control type | Test method | Cut-off | TP | FP | FN | TN |
|------------|------|---------------|--------------|--------------|-------------|---------|----|----|----|----|
| Bind       | 2021 | Cytology + Histology | 27/91        | Benign gallbladder lesions | CMIA    | 252.3 U/mL | 27 | 1  | 0  | 90 |
| Chen       | 2019 | Pathology     | 319/555      | Benign gallbladder lesions | NR      | 37 kU/L | 124 | 34 | 195 | 521 |
| Chen       | 2020 | Pathology     | 103/203      | Benign gallbladder lesions | NR      | 26.52 U/mL | 66 | 12 | 37 | 191 |
| Gao        | 2017 | Pathology     | 30/50        | Benign gallbladder diseases | ECLIA    | 37 U/mL | 15 | 0  | 15 | 50 |
| Gong       | 2013 | Pathology     | 40/80        | Benign gallbladder lesions, Healthy individuals | ECLIA  | 39 U/mL | 35 | 6  | 5  | 74 |
| Huang      | 2014 | Pathology     | 32/52        | Benign bileal diseases | CMIA    | 144.1 U/mL | 17 | 6  | 15 | 46 |
| Kong       | 2021 | Pathology     | 61/61        | Benign gallbladder lesions | NR      | 26.6 U/mL | 40 | 6  | 21 | 55 |
| Li         | 2017 | NR            | 45/45        | Cholecystitis | ECL      | NR      | 38 | 11 | 7  | 34 |
| Liu        | 2000 | Imaging or Pathology | 32/50        | Healthy individuals | ELISA   | 31.5 U/mL | 25 | 2  | 7  | 48 |
| Liu        | 2009 | Pathology     | 18/348       | Benign biliary diseases | CL      | 37 U/mL | 12 | 59 | 6  | 289 |
| Liu        | 2019 | Pathology     | 30/60        | Benign gallbladder lesions, Healthy individuals | CLIA   | 114.31 U/mL | 19 | 5  | 11 | 55 |
| Song       | 2017 | Histopathology | 80/80        | Benign gallbladder diseases | CLIA   | 35 KU/L | 45 | 21 | 35 | 59 |
| Su         | 2008 | Pathology     | 30/50        | Gallbladder polyps | ECL     | 37 U/mL | 23 | 9  | 7  | 41 |
| Ueta       | 2021 | Surgery or EUS-FNA | 50/69        | Benign gallbladder diseases, Healthy individuals | ECLIA | 35 KU/L | 42 | 1  | 22 | 74 |
| Wang       | 2006 | Pathology     | 20/20        | Benign gallbladder lesions | CL      | 37 μU/mL | 16 | 1  | 4  | 19 |
| Wang       | 2016 | Pathology     | 100/131      | Benign gallbladder diseases | ECLIA  | 27 U/mL | 32 | 8  | 68 | 123 |
| Wang       | 2014 | Imaging + pathology | 40/80        | Benign gallbladder lesions, Healthy individuals | ECLIA  | 37 U/mL | 33 | 6  | 7  | 74 |
| Wang       | 2016 | Pathology     | 76/98        | Healthy individuals | ECLIA   | 27 U/mL | 42 | 1  | 34 | 97 |
| Wang       | 2014 | Pathology     | 78/78        | Benign gallbladder diseases | ECLIA  | 39 U/mL | 56 | 4  | 22 | 74 |
| Wu         | 2004 | Pathology or Imaging | 50/40        | Gallbladder polyps | TRFMA   | 37 KIU/L | 32 | 6  | 18 | 34 |
| Wu         | 2016 | NR            | 31/33        | Benign gallbladder lesions | CLIA  | 39.0 U/mL | 23 | 2  | 8  | 31 |
| Yu         | 2014 | Surgery       | 95/94        | Benign gallbladder diseases | ELISA  | 200 μg/mL | 68 | 2  | 27 | 92 |
| Yun        | 2010 | Pathology     | 27           | Healthy individuals | ECL     | NR      | 22 | 4  | 5  | 33 |
| Zhang      | 2019 | Histopathology | 30/60        | Cholecystitis, Healthy individuals | CLIA  | 40 U/mL | 18 | 17 | 12 | 43 |
| Zhang      | 2016 | Histopathology | 35/50        | Benign gallbladder diseases | IRMA   | 39 U/mL | 26 | 3  | 9  | 47 |
| Zhou       | 2014 | Laboratory tests + imaging + Histopathology | 18/19        | Benign gallbladder diseases | CLIA   | 37 U/mL | 14 | 5  | 4  | 14 |
| Zhu        | 2003 | Pathology     | 19/205       | Benign gallbladder diseases | IRMA   | 37 U/mL | 18 | 56 | 1  | 149 |

CL: chemiluminescence; CLIA: chemiluminescent immunoassay; CMIA: chemiluminescent microparticle immunoassay; ECL: electrochemiluminescence; ECLIA: electrochemiluminescent immunoassay; EUS-FNA: endoscopic ultrasonography-guided fine needle aspiration; FN: false negative; FP: false positive; GBC: gallbladder cancer; IRMA: immunoradiometric assay measurement; NR: not reported; Pts: patients; Ref: reference; TN: true negative; TP: true positive; TRFMA: time-resolved fluorinated assay.
studied tumor marker primarily applied for the management of pancreato-biliary tumors. However, the current use of CA19-9 for GBC is limited. At present its stand-alone application has not been recommended by international guidelines, such as the National Comprehensive Cancer Network Guidelines for Hepatobiliary Cancers,41 although previous reports have highlighted the potential of serum CA19-9 in predicting resectability, tumor burden, and recurrence in patients with GBC.7–9,42 Of note, the diagnostic role of CA19-9 has also yet to be explained in GBC. Therefore, the present study—the first meta-analysis to the best of knowledge—was undertaken to seek evidence for the potential application of serum CA19-9 for the detection of GBC through a collation of the available literature.

A total of 27 studies involving 1516 GBC cases and 2784 non-GBC controls contributed data to this meta-analysis. Consistent with epidemiological trends related to geographic locations, that a greater than average incidence and mortality rate of GBC is noted in Asia regions1,43 and China is a high-GBC risk country where GBC cases account for nearly one-quarter of global cases,44 all the included studies came from Asian countries and most were carried out in China. Pooled

---

**Figure 1.** Forest plots for (a) sensitivity and (b) SPE of CA19-9 in diagnosis of gallbladder cancer.

**Figure 2.** Summary receiving operator characteristics curves (SROC) of CA19-9 in the diagnosis of gallbladder cancer. Included studies are arranged in alphabetical order (A–Z) by the first letter of the first author: (1) Bind et al.’s study; (27) Zhu et al.’s study.

AUC: area under the curve; SENS: sensitivity; SPEC: specificity.
results showed that serum CA19-9 maintained moderate sensitivity (0.70) and good SPE (0.92) in the diagnosis of GBC, with an AUC of 0.89 considered to be in the moderate range. The combined PLR was 8.30, implying that a patient with GBC is 8.3-fold more likely to have positive CA19-9 result than that without GBC. Likewise, the pooled NLR of 0.33 indicated that the probability of true negatives increases by 67% when the CA19-9 testing was negative. An overall DOR of 25.12 supported CA19-9 as a helpful tumor marker for the diagnosis of GBC based on the fact that a DOR of 1 suggested an inability to differentiate GBC from controls and that higher values indicated better discriminatory test performance.13 Elevated CA19-9 in serum was previously reported to have a moderate accuracy in the diagnosis of biliary tract carcinoma containing GBC and cholangiocarcinoma (CCA).45 Liang et al. undertook a meta-analysis of 31 eligible studies to evaluate the role of serum CA19-9 in diagnosing CCA and showed that the SEN, SPE, PLR, NLR, DOR, AUC were 0.72, 0.84, 4.93, 0.35, 15.10, and 0.84, respectively.46 The results in the present study shared similarity with their results in terms of the relatively lower SEN and moderate AUC. Notably enough, as demonstrated by the subgroup and meta-regression analysis, both the combined sensitivity and SPE were significantly lower in the subgroup where there was a majority of subjects with benign gallbladder diseases, such as cholecystitis and gallbladder polyps (sensitivity: 0.69, SPE: 0.91), than in the subgroup of cohorts mixed with many healthy individuals (sensitivity: 0.72, SPE: 0.93) (p < 0.05). This represents a challenge for the interpretation of elevated CA19-9 levels in benign biliary conditions. Given the overall moderate accuracy of CA19-9 for detecting GBC and its declined potential in differentiating malignant and benign gallbladder lesions, more attention should be paid to the factors that could affect its performance to improve the diagnostic accuracy of CA19-9 and validate its universal use in the future. Of note is the impossibility of CA19-9 detection in subjects with a fucosyltransferase deficiency, which might increase the proportion of FNs and result in a lower sensitivity. Based on that simultaneous detection of CA 19-9 (sialyl Lewis-a), and its counterpart, namely disialyl Lewis-a, remains a useful technique to overcome this limitation.7,45 Also, the presence of obstructive jaundice could serve as a factor at play as it is likely to increase the serum CA19-9 levels in both malignant and benign biliary diseases.47 In this case, repeated measurement of CA19-9 should be carried out after the successful treatment of jaundice, using the same positive criteria set before, to exclude the interference of jaundice. As inflammation or

Table 2. Results of the sensitivity analysis.

| Included studies | Sensitivity (95% CI) | I^2 | p  | SPE (95% CI) | I^2 | p  |
|------------------|----------------------|-----|----|--------------|-----|----|
| All              | 0.70 (0.63–0.76)     | 87.42 | 0.00 | 0.92 (0.88–0.94) | 91.39 | 0.00 |
| **Removal**      |                      |      |    |              |      |    |
| (1) Bind et al.  | 0.68 (0.62–0.74)     | 86.12 | 0.00 | 0.91 (0.87–0.93) | 90.46 | 0.00 |
| (27) Zhu et al.  | 0.69 (0.62–0.75)     | 87.55 | 0.00 | 0.92 (0.89–0.94) | 88.75 | 0.00 |
| (16) Wang Q. et al. | 0.71 (0.65–0.77) | 85.38 | 0.00 | 0.91 (0.88–0.94) | 91.24 | 0.00 |
| (1) + (16)       | 0.69 (0.64–0.75)     | 83.67 | 0.00 | 0.91 (0.87–0.93) | 90.26 | 0.00 |
| (16) + (27)      | 0.70 (0.64–0.75)     | 85.48 | 0.00 | 0.92 (0.89–0.94) | 88.58 | 0.00 |
| (1) + (16) + (27)| 0.68 (0.63–0.73)     | 83.57 | 0.00 | 0.91 (0.88–0.94) | 87.26 | 0.00 |

Included studies are arranged in alphabetical order (A–Z) by the first letter of the first author during sensitivity analysis: (1) Bind et al.’s study; (16) Wang Q et al.’s study; (27) Zhu et al.’s study.

CI: confidence interval; No.: number; SPE: specificity.

Table 3. Results of subgroup and meta-regression analysis.

| Subgroups                  | No. of study | Sensitivity (95% CI) | p       | SPE (95% CI) | p       |
|----------------------------|--------------|----------------------|---------|--------------|---------|
| **Control type**           |              |                      |         |              |         |
| Benign biliary diseasesa   | 19           | 0.69 (0.61–0.77)     | 0.04*   | 0.91 (0.87–0.95) | 0.00*   |
| Mixed cohortsb             | 8            | 0.72 (0.61–0.83)     | 0.93 (0.88–0.97) |         |
| **Sample size**            |              |                      |         |              |         |
| <100                       | 12           | 0.72 (0.63–0.81)     | 0.12    | 0.90 (0.85–0.95) | 0.00*   |
| ≥100                       | 15           | 0.68 (0.59–0.77)     | 0.93 (0.90–0.96) |         |

*a*statistical significance. 
*aMost of these were benign gallbladder diseases such as cholecystitis, gallbladder polyps, gallbladder stones.
*bMixed cohorts included (1) only healthy controls; (2) healthy controls + benign gallbladder diseases.

CI: confidence interval; No.: number; SPE: specificity.
obstruction of biliary tracts can contribute to increased CA19-9 levels, significant elevation was also possibly observed in benign biliary diseases.\textsuperscript{7,45} It is therefore plausible for CA19-9 to act as a useful adjunct to other diagnostic tools, such as imaging tests, rather than as a sole candidate marker for the detection of GBC. It is also recommended that a combined assay of CA19-9 with other serological markers, such as CEA, CA125, CA153, and CA142, be conducted.\textsuperscript{19,22,26,31,32,34,37,38} For example, Wang et al.\textsuperscript{10} demonstrated that the combination of CA199, CA242, and CA125 showed the highest accuracy compared with each of these markers alone. When the positive result was made based on one positive of any parameter within the combination, the sensitivity and SPE for GBC diagnosis were 91\% and 91.4\%, respectively. More studies are needed in the future to explore the clinical value of CA19-9 combined with other markers to establish a useful marker-based diagnostic recommendation in the management of GBC. In addition, serum CA19-9 may process a better SPE in a larger-size group (≥100, SPE: 0.93) than in a smaller-size group (<100, SPE: 0.90) (p < 0.05), indicating that serum CA19-9 may reduce the misdiagnosis of GBC in a large scale. Hence, future research can take this factor into consideration and more larger-scale studies are needed to further confirm its influence.

Notably, the threshold effect acted as the main source of high heterogeneity in the overall results. A threshold effect analysis was performed, which suggested the presence of a threshold effect. A traditional cut-off value for serum CA19-9 is 37 U/mL, which is typically used for PC, with normal values ranging from 0 to 37 U/mL.\textsuperscript{7,45} This cut-off value was also applied and studied for GBC in several of the included papers, yet an optimal threshold for this condition still needs to be established. More importantly, variability in the methods for measuring serum CA19-9 levels and test assays from various manufacturers could also lead to a different range of thresholds among the included studies. It is implied that for accurate interpretation further studies should investigate the optimal cut-off level of CA19-9 in diagnosis of GBC; standardization of measuring methods for CA19-9 detection is of significance and the same assay should be used to consistently monitor the suspected GBC cases whose CA19-9 levels show an uptrend. Additionally, sensitivity analyses were subsequently performed to further explore the potential source of heterogeneity. Despite the recognition of outlying studies, their removal did not make a significant impact on the heterogeneity and the general trend of overall results, which to a certain extent indicated the robustness of outcomes.

Limitations
The first limitation is the existence of considerable heterogeneity in the analysis due to the threshold effect. Combined results should be viewed with caution as the diagnostic values will be different between included studies that used varied thresholds and more attention should be paid to the AUC figure and the SROC curve (Figure 2), which displayed an overall evaluation of diagnostic performance of CA19-9 among individual studies. Additionally, the included papers were limited by their case-control design in a retrospective way (there was only one prospective study). They enrolled subjects with confirmed GBC cases for their analyses, but an ideal diagnostic study should enroll a proportion of patients with suspected GBC to reduce the risk of bias in terms of patient selection. Therefore, future research is required to avoid this issue and provide more robust evidence. Lastly, the eligible studies were written in English or Chinese and were conducted in the same region (Asia). There is limited evidence for this study to investigate whether or not geographic location and ethnicity influences the diagnostic accuracy of serum CA19-9; therefore, additional studies are encouraged for further confirmation.

Conclusions
This meta-analysis provides evidence for serum CA19-9 to be a candidate marker for the detection of GBC. CA19-9 in serum exhibits moderate sensitivity and good SPE, with overall diagnostic performance in the moderate range. Its diagnostic performance can be influenced when it is used for distinguishing GBC from benign biliary diseases, and the sample size might also affect its value. CA19-9 in serum is recommended to be used in combination with other diagnostic tools in the diagnosis of GBC. Results should be interpreted with caution as there significant heterogeneity exists, mainly owing to the threshold effect, which might be due to the varied cut-off values in the included studies. More well-designed studies are needed to establish an optimal cut-off value for CA19-9 in the diagnosis of GBC and to further validate its clinical application.

Acknowledgments
Not applicable.

Funding
The author declares that no funding support was received.

Declaration of conflicting interests
The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID ID
Xiaolei Zhou https://orcid.org/0000-0002-2699-9647

Supplemental material
Supplemental material for this article is available online.
References

1. Ferlay J, Ervik M, Lam F, et al. (2020). Global cancer observatory: Cancer today. Lyon, France: International Agency for Research on Cancer. Available from: https://gco.iarc.fr/today/data/factsheets/cancers/12-Gallbladder-fact-sheet.pdf.

2. Hundal R and Shaffer EA. Gallbladder cancer: epidemiology and outcome. *Clin Epidemiol* 2014; 6: 99–109.

3. Garcia P, Lamarca A, Diaz J, et al. Current and new biomarkers for early detection, prognostic stratification, and management of gallbladder cancer patients. *Cancers (Basel)* 2020; 12: 367014.

4. Rizzo A and Brandi G. Pitfalls, challenges, and updates in adjuvant systemic treatment for resected biliary tract cancer. *Expert Rev Gastroenterol Hepatol* 2021; 15: 547–554.

5. Rizzo A and Brandi G. BILCAP trial and adjuvant capcetabine in resectable biliary tract cancer: reflections on a standard of care. *Expert Rev Gastroenterol Hepatol* 2021; 15: 483–485.

6. Huang Z and Liu F. Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a meta-analysis. *Tumour Biol* 2014; 35: 7459–7465.

7. Lee T, Teng TZJ and Shelat VG. Carbohydrate antigen 19-9 - tumor marker: past, present, and future. *World J Gastrointest Surg* 2020; 12: 468–490.

8. Agraval S, Lawrence A and Saxena R. Does CA 19-9 have prognostic relevance in gallbladder carcinoma (GBC)? *J Gastrointest Cancer* 2018; 49: 144–149.

9. Liu F, Wang JK, Ma WJ, et al. Clinical value of preoperative CA19-9 levels in evaluating resectability of gallbladder carcinoma. *ANZ J Surg* 2019; 89: E76–E80.

10. Wang YF, Feng FL, Zhao XH, et al. Combined detection of tumor markers for diagnosis and prognosis of gallbladder cancer. *World J Gastroenterol* 2014; 20: 4085–4092.

11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *PloS Med* 2021: e1003583.

12. 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: 529–536.

13. Eusebi P. Diagnostic accuracy measures. *Cerebrovasc Dis* 2013; 36: 267–272.

14. Macaskill P, Gatsonis C, Deeks JJ, et al. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C (eds), Cochrane Handbook for systematic reviews of diagnostic test accuracy version 1.0. The cochrane collaboration, 2010. Available from: http://srdata.cochrane.org/.

15. Bind MK, Mishra RR, Kumar V, et al. Serum CA 19-9 and CA 125 as a diagnostic marker in carcinoma of gallbladder. *Indian J Pathol Microbiol* 2021; 64: 65–68.

16. Chen M, Cao J, Bai Y, et al. Development and validation of a nomogram for early detection of malignant gallbladder lesions. *Clin Transl Gastroenterol* 2019; 10: e00998.

17. Chen Z, Liu Z, Zhang Y, et al. Combination of CA19-9 and the neutrophil-to-lymphocyte ratio for the differential diagnosis of gallbladder carcinoma. *Cancer Manag Res* 2020; 12: 4475–4482.

18. Gao L, Chen S and Wang G. Effect of ultrasound combined with determination of CA19-9 and CEA on the accuracy of early diagnosis of gallbladder cancer. *Chin J Surg Integr Tradit West Med (Chin)* 2017; 23: 239–242.

19. Gong P, Yao C and Zhao X. Diagnostic value of combined detection of serum CA199, CA125, CEA, and CA153 for gallbladder carcinoma. *Shandong Med J* 2013; 53: 58–59.

20. Huang L, Chen W, Liang PW, et al. Serum CYFRA 21–1 in biliary tract cancers: a reliable biomarker for gallbladder carcinoma and intrahepatic cholangiocarcinoma. *Dig Dis Sci* 2014; 60: 1273–1283.

21. Kong WH, Zhang L, An R, et al. Diagnostic value of serum D-dimer for detection of gallbladder carcinoma. *Cancer Manag Res* 2021; 13: 2549–2556.

22. Li Z. Exploring the clinical value of combined detection of CEA, CA19-9, and CA125 in the diagnosis of gallbladder carcinoma. *Orien Diet Ther Health Care (Chin)* 2017; 10: 249.

23. Liu K and Ge X. Clinical analysis of 32 cases of gallbladder carcinoma detected by tumor marker CA19-9. *J Bengbu Med Coll (Chin)* 2000; 25: 463–464.

24. Liu KM. Application of serum carbohydrate antigen 19-9 and carcinoembryonic antigen in diagnosis of biliary carcinomas. *J Chin Prescription Drug (Chin)* 2019; 17: 129–130.

25. Liu S and Yuan H. The further evaluation of the clinical significance of serum CA199 and its diagnostic value of hepatobiliary and pancreatic diseases. *J Med Theor Pract (Chin)* 2009; 22: 1417–1419.

26. Song Z, Liao Y and Yuan J. Clinical significance of combined detection of serum CA199 and CA125 in patients with gallbladder carcinoma. *Psychol Doct (Chin)* 2017; 23: 55–57.

27. Su M and Zhao X. Diagnostic value of detection of CA19-9, CEA, AFP, and CA125 in gallbladder carcinoma patients. *Med J Chin Armed Police Forces (Chin)* 2008; 19: 906–908.

28. Ueta E, Tsutsumi K, Kato H, et al. Extracellular vesicle-shuttled miRNAs as a diagnostic and prognostic biomarker and their potential roles in gallbladder cancer patients. *Sci Rep* 2021; 11: 13.

29. Wang Q, Pan S and Chen Y. Clinical significance of serum carbohydrate antigen 19-9 level in gallbladder cancer patients. *Chin J Clin Med (Chin)* 2006; 13: 591–592.

30. Wang Q, Qi J and Hu L. Clinical significance of combined detection of serum tumor markers in diagnosis of gallbladder carcinoma. *Int J Lab Med (Chin)* 2016; 37: 106–107.

31. Wang S and Zhang D. Three tumor markers combined detection in diagnosis of gallbladder cancers. *Chin J Surg Integr Tradit West Med (Chin)* 2014; 20: 485–487.

32. Wang X, Zheng X and Li G. Diagnostic value of combined detection of tumor markers in diagnosis of gallbladder carcinoma. *J Clin Med Literature (Chin)* 2016; 3: 9606.

33. Wu L, Huang Q, Fu B, et al. Role of combined testing of serum tumor markers in diagnosis of gallbladder carcinoma. *J Hepatopancreatobiliary Surg (Chin)* 2004; 16: 32–34.

34. Wu Y. Efficacy of serum CA199, CA242 and CA125 in diagnosis of gallbladder carcinoma. *Chin Med Devices (Chin)* 2016; 31: 139–140.

35. Yu F, Wang Y, Wu X, et al. Diagnostic value of detection of S100A11 in gallbladder carcinoma patients. *Chin J Mod Med (Chin)* 2014; 24: 37–41.

36. Yun W and Wang H. The diagnosing value of tumor markers in serum for five digestive system cancers. *J Qinghai Med Coll (Chin)* 2010; 31: 92–95.
37. Zhang Y and Yang Q. Diagnostic value of combined detection of serum CA125, CA199, and CEA in diagnosis of gallbladder carcinoma. *Mod Dig Interv (Chin)* 2019; 24: 1465–1467.

38. Zhang YD. Serum CA199, CA242, CA125 joint clinical value in the diagnosis of gallbladder carcinoma detection. *J Hepatopancreatobiliary Surg (Chin)* 2016; 24: 106–110.

39. Zhou J, Chen H and Chen Y. Diagnostic value of five serum tumor markers in digestive system tumors and benign diseases by microparticle enzyme immunoassay. *Chin J Health Lab Tec (Chin)* 2014; 24: 972–975.

40. Zhu C, Shen J, Huang G, et al. The clinical value of CA-242 and CA19-9 immunoassay in differential diagnosis of benign and malignant gallbladder disease and the influence of hyperbilirubinemia. *Nuclear Tech (Chin)* 2003; 26: 367–370.

41. Benson AB, D’Angelica MI, Abbott DE, et al. Hepatobiliary cancers, version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *JNCCN* 2021; 19: 541–565.

42. Sachan A, Saluja SS, Nekarakanti PK, et al. Raised CA19-9 and CEA have prognostic relevance in gallbladder carcinoma. *BMC Cancer* 2020; 20: 826.

43. Schmidt MA, Marcano-Bonilla L and Roberts L.R. Gallbladder cancer: epidemiology and genetic risk associations. *Chin Clin Oncol* 2019; 8: 31.

44. Ren T, Li Y, Zhang X, et al. Protocol for a gallbladder cancer registry study in China: the Chinese research group of gallbladder cancer (CRGCG) study. *BMJ Open* 2021; 11: e038634.

45. Tsen A, Barbara M and Rosenkranz L. Dilemma of elevated CA19-9 in biliary pathology. *Pancreatology* 2018; 18: 862–867.

46. Liang B, Zhong L, He Q, et al. Diagnostic accuracy of serum CA19-9 in patients with cholangiocarcinoma: a systematic review and meta-analysis. *Med Sci Monit* 2015; 21: 3555–3563.

47. Lin MS, Huang JX and Yu H. Elevated serum level of carbohydrate antigen 19-9 in benign biliary stricture diseases can reduce its value as a tumor marker. *Int J Clin Exp Med* 2014; 7: 744–750.