Association between gallstones and the risk of biliary tract cancer:
a systematic review and meta-analysis

Dan Huang\textsuperscript{a,e,*}, Hyundeok Joo\textsuperscript{d,*}, Nan Song\textsuperscript{b,d}, Sooyoung Cho\textsuperscript{a}, Woosung Kim\textsuperscript{a}, Aesun Shin\textsuperscript{a,b}

\textsuperscript{a}Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

\textsuperscript{b}Cancer Research Institute, Seoul National University, Seoul, Republic of Korea

\textsuperscript{c}Department of Epidemiology and Cancer Control, St. Jude Children’s Research Hospital, Memphis, TN, USA

\textsuperscript{d}Seoul National University College of Medicine, Seoul, Republic of Korea

\textsuperscript{e}Division of Cancer Control & Policy, National Cancer Control Institute, National Cancer Center, Goyang, Republic of Korea

\textsuperscript{*}D. Huang and H. Joo are co-first authors and contributed equally to this work

Corresponding author

Aesun Shin, MD, PhD

Department of Preventive Medicine, College of Medicine, Seoul National University

103 Daehak-ro, Jongno-gu, Seoul, Korea, 03080

Tel: +82-2-740-8331 / Fax: +82-2-747-4830 / E-mail: shinaesun@snu.ac.kr
Association between gallstones and the risk of biliary tract cancer:
a systematic review and meta-analysis
Abstract

Objectives: Biliary tract cancers (BTCs) are rare but highly fatal. Although the etiology of BTC is poorly understood, gallstones are proposed to be an imperative risk factor. We conducted a systematic review and meta-analysis to examine the associations between gallstone characteristics and the risk of BTC.

Methods: We searched the Medline, Embase, and Cochrane Central databases and systematically reviewed cohort and case-control studies published before April 9, 2018. All the included studies reported appropriate risk estimates and their confidence intervals (CIs) on the association between the presence, size, number, or duration of gallstones and the risk of BTC, including gallbladder cancer (GBC), extrahepatic bile duct cancer (EBDC), and ampulla of Vater cancer (AOVC). Summary odds ratios (ORs) and their 95% CIs were calculated using a random-effects model meta-analysis. Subgroup analyses were conducted to inspect the source of potential heterogeneity, and Egger's test was performed to assess publication bias.

Results: A total of seven cohort studies and 23 case-control studies in the Asian, European, and American populations were included. The presence of gallstones was associated with an increased risk of BTC (OR=4.38; 95% CI=3.23-5.93; I²=91.2%), GBC (OR=7.26; 95% CI=4.33-12.18), EBDC (OR=3.17; 95% CI=2.24-4.50), and AOVC (OR=3.28; 95% CI=1.33-8.11). Gallstone size (>1 cm vs. <1 cm; OR=1.88, 95% CI=1.10-3.22) was significantly associated with the risk of GBC.

Conclusions: This systematic review and meta-analysis indicates that gallstone characteristics, such as presence, size, and number, are associated with an increased risk of BTC. However, the study has limitations that significantly high heterogeneities were
present in the meta-analyses.

Keywords

Gallstones; Biliary Tract Neoplasms; Gallbladder Neoplasms; Bile Duct Neoplasms; Cholangiocarcinoma; Ampulla of Vater.

Abbreviations

BTC, biliary tract cancer; GBC, gallbladder cancer; EBDC, extrahepatic bile duct cancer; AOVC, ampulla of Vater cancer; OR, odds ratio; CI, confidence interval; NOS, Newcastle-Ottawa Scale; EHC, extrahepatic cholangiocarcinoma; BDC, bile duct cancer; CCA, cholangiocarcinoma; IBDC, intrahepatic bile duct cancer
1. Introduction

Biliary tract cancer (BTC) is a relatively uncommon cancer in most parts of the world [1], yet it is classified as a major cancer based on its incidence in certain countries such as the Republic of Korea and India (New Delhi) [2, 3]. The increasing incidence of BTC is observed in all three biliary tract subsites, specifically gallbladder cancer (GBC), extrahepatic bile duct cancer (EBDC), and ampulla of Vater cancer (AOVC), especially within high-risk areas [4, 5]. The prognosis of BTC is generally poor, and the estimated 5-year survival rate is only approximately 5% [6]. In terms of the treatment option for BTC, although surgery can be curative, a small percentage of patients are candidates for surgery [7]. This is because the high proportion of patients is diagnosed at a late stage of the disease [7]. To improve the survival rate due to this poor prognosis, early detection of the disease through identifying risk factors is important.

Gallstones, concretions formed in the biliary tract, have been suggested as one of the important risk factors for BTC [5]. The carcinogenic mechanisms of BTC are poorly understood, but they may involve inflammatory changes near stones [8]. BTC could arise as a result of chronic inflammation associated with gallstones continuously irritating the gallbladder and bile duct [9]. While gallstones are common conditions among the population [10], BTC rarely occurs, and most people with gallstones never end up developing cancer [11, 12]. However, a significant number of BTC patients have gallstones [13], which leaves room for further investigations on the association between gallstones and the risk of BTC. One study attempted a systematic review [14], but it examined the literature on the association between benign gallbladder disease (the
broader term to represent gallstones) and the risk of BTC. There is scarcity of reviews focusing on the relationship between gallstones and the risk of BTC.

We conducted a systematic review and meta-analysis of published cohort and case-control studies on associations between gallstone characteristics and the risk of BTC. This study aimed to update the latest studies through systematic review and to provide a better description of the association of gallstones with the risk of BTC, encompassing its known subtype GBC, EBDC, and AOV [15], whereas intentionally excluding intrahepatic bile duct cancer (IBDC) or intrahepatic cholangiocarcinoma (IHC). The objective is that with the vast populations combined from various studies (case-control and cohort studies) throughout the world, finding out to how extent patients with gallstones more likely to develop BTC and each of its subtypes than the hospital or community-based control groups. In this study, gallstones were characterized by the presence, size, number, and duration, and detailed subgroup analyses were also performed stratified by the study design, sex, geographic areas, study period, measurement of exposure, study quality score, and adjustment of confounders.

2. Materials and Methods

The study protocol followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines.

2.1. Data sources and searches

The first and second reviewers (DH. and HJ.) searched the PubMed, Embase and Cochrane Library databases for epidemiological studies with the following keywords:
The MeSH (medical subject headings) terms were used for the PubMed search, and Emtree explode terms were used for Embase search when available. The last search was conducted on April 9, 2018. The language was restricted to English in PubMed and Embase but Cochrane library databases. In terms of publication status, our search was confined only to published human studies. Papers published before April 9, 2018, were reviewed. Duplicates were excluded, and additional papers obtained by manually searching the references of the selected articles were included.

2.2. Study selection

The inclusion criteria for eligible studies were as follows: (1) cohort study or case-control study on the association between gallstones and the risk of BTC (GBC, EBDC, or AOVC); (2) gallstones (presence, size, number, or duration) as the exposure of interest; (3) primary outcome was the occurrence of BTC (GBC, EBDC, or AOVC); and (4) studies reported risk estimates (rate ratio [RR], odds ratio [OR], or hazard ratio [HR]) and their 95% confidence intervals (CIs). Studies were excluded if any of the following criteria were met: (1) nonhuman studies; (2) nonobservational studies or observational studies without analytical epidemiologic approach; (3) irrelevant exposure or outcome variable (hepatolithiasis or intrahepatic cholangiocarcinoma); (4) in case of duplication
or unobtainable abstract/full-text; (5) no risk estimate is reported or could be calculated by the given information.

2.3. Data extraction

The first and third reviewers (DH. and WK.) (under the supervision of AS.) independently screened the titles and abstracts of studies that met the inclusion criteria. The full texts were reviewed by two independent reviewers (DH. and HJ.) and the supervisor (AS.).

The two independent reviewers (DH. and HJ.) extracted data using a standardized extraction form. When discrepancies arose, a fourth investigator (NS.) made the final decision for study eligibility and data extraction. The relevant data included the last name of the first author, publication year, study country, study design (cohort or case-control study), study period, sex, sample size (number of cohorts and incident cases for cohort studies or number of cases and controls for case-control studies), exposure variables (presence, size, number, and duration of gallstones), measurement of exposure (with or without imaging study), outcome variables (occurrence of GBC, EBDC, AOVC), duration of follow-up for cohort studies, adjustment variables in the statistical analysis, and risk estimates, such as OR, RR and HR with corresponding 95% CIs.

2.4. Quality Assessment

The quality assessment data were extracted using the Newcastle-Ottawa Scale (NOS), which contains 9 items with 8 items receiving 1 point and 1 item accounting for 2 points,
leading to a maximum of 10 points \[16\]. A quality score equal to or greater than the median value was judged as high quality.

**2.5. Statistical analysis**

In this study, the summary risk estimates and their corresponding 95% CIs were calculated using a random effects model \[17\]. Selected studies reported different types of risk estimates, such as ORs, RRs, and HRs. RRs and HRs were treated as equivalent to ORs. We compared gallstone characteristics as follows: presence (present vs. absent), size (≥1 cm vs. <1 cm, ≥2 cm vs. <2 cm), numbers (>1 vs. 1), and duration. In the studies reporting multiple risk estimates according to the subsites of BTC (GBC, EBDC, and/or AOVC), the pooled risk estimates and their corresponding 95% CIs that were adequate for meta-analysis were taken as representative risk estimates.

Statistical heterogeneity across studies was appraised using $I^2$ statistics and Chi-square-based Q-tests. $I^2$ values of 25%, 50%, and 75% indicated low, moderate, and high heterogeneity, respectively \[18\]. For the Q statistics, a $p$-value <0.10 was considered statistically significant for heterogeneity. To perform subgroup analyses, we stratified studies by the study design, sex, geographic area (Asia and non-Asia), study period (before, around and after 2000; around 2000 refers to the studies where the starting point is before 2000 but the ending point is after 2000), measurement of exposure, study quality according to NOS, and whether adjusted for confounders (such as age, sex, comorbidity, lifestyle factors, education, and/or geographic areas).

Sensitivity analyses \[19\] were conducted by sequentially excluding one study in one turn to evaluate the influence of individual studies on the stability of pooled results. Forest plots were used to present graphical results. Publication bias was investigated
through funnel plots [20] with Egger's test [21], and $p < 0.01$ indicated statistical significance. All statistical analyses were performed using STATA, version 15.0 (STATA, College Station, TX, U.S.). A two-tailed $p$-value $<0.05$ was considered statistically significant.

3. Results

3.1. Study selection and characteristics

Fig. 1 shows the process of study selection for the meta-analysis. Initially, we retrieved a total of 5,005 articles, including 1,941 from Medline, 3,027 from Embase, and 37 from the Cochrane Library. We excluded 1,082 duplicate studies. Based on reviewing the titles and abstracts, 3,751 studies were also excluded for various reasons: animal studies (n=2); non-observational studies (n=786); irrelevant exposures or outcomes (n=2,957); and no abstract or full text (n=6). We reviewed the full texts of the remaining 172 studies and excluded articles with irrelevant exposures or outcomes, insufficient data for the meta-analysis, or other exclusion criteria, thus resulting in 27 eligible articles. Additionally, three relevant studies were included by searching the reference lists of the eligible articles. Thus, a total of 30 epidemiological studies, including seven cohort studies and 23 case-control studies, were included in this meta-analysis.

The characteristics of 7 cohort studies and 23 case-control studies are shown in Table 1. Of these studies, 16 studies were conducted in Asia, eight studies were conducted in America, and six studies were conducted in Europe. Associations between gallstones and the risk of GBC were investigated in 18 studies, where gallstones were characterized as their presence in 14 studies, their size in four studies, and their number
in two studies. In terms of the risk of EBDC, with the concept of EBDC embracing extrahepatic cholangiocarcinoma (EHC), bile duct cancer (BDC), and cholangiocarcinoma (CCA), there were 17 studies that examined the association between gallstone presence with the risk of BTC. Among five studies on AOVC, all studies reported an association between gallstone presence and the risk of cancer.

3.2. Gallstones and the risk of BTC

A total of 26 studies presented associations between the presence of gallstones and the risk of BTC (Fig. 2). Among these studies, only two studies referred to BTC specifically, and the remaining 24 studies described the risk estimates according to the subsites of BTC (GBC, EBDC, and/or AOVC).

We identified seven cohort studies and 19 case-control studies that presented associations between the presence of gallstones and the risk of BTC. When we examined the results stratified by the study design, a statistically significant positive association was shown in both case-control studies (OR=5.04; 95% CI 3.36-7.56; I²=90.5%; p<0.001) and cohort studies (OR=3.17; 95% CI 2.28-4.39; I²=79.0%; p<0.001). The pooled risk estimate was also statistically significant (OR=4.38; 95% CI 3.25-5.93), with high heterogeneity across the studies (I²=91.2%; p<0.001).

In subgroup meta-analyses, all results showed statistical significance, regardless of sex, geographic area, study period, measurement of exposure, study quality, and adjustment for confounders, as shown in Table 2. The magnitudes of associations were larger in females (OR= 4.26; 95% CI 2.75-6.59; I²=84.5%; p<0.001) than in males, larger in Asia (OR= 5.25, 95% CI 3.50-7.86; I²=82.4%; p<0.001) than outside of Asia, larger in studies conducted before 2000 (OR=5.39, CI 2.57-11.34; I²=95.5%; p<0.001).
than in studies conducted around and after 2000, larger in studies with imaging study (OR=7.09; 95% CI 3.87-12.98; \( I^2=64.5\% \); \( p=0.004 \)) than in studies without imaging study, and larger in low-quality studies (OR=4.81, 95% CI 2.87-8.05; \( I^2=94.9\% \); \( p<0.001 \)) than in high-quality studies. The meta-analysis indicated that the association became weaker after adjusting for age, sex, and comorbidities. A stronger association than the original result was observed after adjusting for geographic areas, lifestyle factors, and education. However, there were no significant differences between the magnitudes of association under any stratifications.

The heterogeneity varied substantially as the stratification method changed, and the male group studies exhibited the lowest level of heterogeneity (\( I^2=35.8\% \); \( p=0.132 \)) among the subgroup studies including more than two studies.

3.3. Gallstones and the risk of GBC

Among the 20 studies on associations between gallstones and the risk of GBC, 16 studies presented associations between the presence of gallstones and the risk of GBC, as shown in Fig. 3A. A total of 5 cohort studies and 11 case-control studies were included in the meta-analysis of this subsite cancer. When we observed the results divided by the study design, statistically significant positive associations were shown in both case-control studies (OR=9.60; 95% CI 4.45-20.70; \( I^2=95.4\% \); \( p<0.001 \)) and cohort studies (OR=4.54; 95% CI 2.62-7.87; \( I^2=72.5\% \); \( p=0.006 \)). The pooled risk estimate including case-control and cohort studies was also statistically significant (OR=7.26; 95% CI 4.33-12.18), with high heterogeneity across the studies (\( I^2=93.6\% \); \( p<0.001 \)).
Meta-analyses were stratified by diverse subgroups, as presented in Table 3. Regardless of the subgroups, all the results of meta-analyses were statistically significant with little differences in the magnitude of risk estimates. However, the differences between the risk estimates according to the subgroup analyses were only statistically significant in the following subgroups: geographic areas of Asia (OR=12.72; 95% CI 6.35-25.46; $I^2=86.2\%$; $p<0.001$) versus non-Asian areas (OR=3.59; 95% CI 2.68-4.81; $I^2=56.0\%$; $p=0.026$), measurement of exposure with imaging study (OR=15.27; 95% CI 7.48-31.18; $I^2=76.9\%$; $p=0.002$) versus without imaging study (OR=4.67; 95% CI 3.29-6.61; $I^2=76.1\%$; $p<0.001$), and adjustment for education (OR=23.80; 95% CI 17.00-33.32) versus the original summary risk estimates (OR=7.26; 95% CI 4.33-12.18; $I^2=93.6\%$; $p<0.001$).

With regard to gallstone characteristics, we found that the risk of GBC was associated with gallstone size (>1 cm vs. <1 cm; OR=1.88, 95% CI=1.10-3.22; $I^2=35.2\%$; $p=0.201$) (>2 cm vs. <2 cm; OR=2.62, 95% CI=0.90-7.60; $I^2=73.8\%$; $p=0.022$) [22-25] and gallstone number (>1 vs. 1; OR=2.10, 95% CI=0.80-5.47; $I^2=63.8\%$; $p=0.096$) [24, 25].
3.4. Gallstones and the risk of EBDC

A total of 17 studies presented the associations between the presence of gallstones and the risk of EBDC in its broadest sense (a concept embracing EBDC, EHC: extrahepatic cholangiocarcinoma, CCA: cholangiocarcinoma, BDC: bile duct cancer), as shown in Fig. 3B. We identified four cohort studies and 13 case-control studies that presented associations between the presence of gallstones and the risk of EBDC. Among the 17 studies, 12 studies reported the risk of EBDC (or EHC), while the remaining six studies accounted for the risk of CCA (or BDC) [26-31], with one study [31] describing the risk of both EBDC and CCA. The summary risk estimate on association between gallstone presence and the risk of cancer was stronger within the studies on EBDC (or EHC) (OR=2.87, 95% CI=2.06-3.99; I²=95.0%; p<0.001) than the studies on CCA (or BDC) (OR=2.12, 95% CI=1.35-3.33; I²=92.7%; p<0.001) without statistical significance.

In the meta-analysis of the comprehensive EBDC, when we obtained the results divided by the study design, a statistically significant positive association was shown in both case-control studies (OR=3.67; 95% CI 2.26-5.95; I²=96.0%; p<0.001) and cohort studies (OR=2.33; 95% CI 2.00-2.72; I²=21.4%; p=0.282). The pooled risk estimate was also statistically significant (OR=3.17; 95% CI 2.24-4.50), with high heterogeneity across the studies (I²=95.2%; p<0.001).

In the subgroup meta-analyses, all results showed statistical significance regardless of sex, geographic area, study period, measurement of exposure, and study quality, as presented in Table 4. However, the differences between the magnitudes of the effect sizes did not have statistical significance in any way of the stratifications.
3.5. Gallstones and the risk of AOVC

A total of five studies presented associations between gallstone characteristics and the risk of AOVC. Among these studies, 1 study reported the duration of the gallstones [27], and all five studies reported on the presence of gallstones [13, 27, 30, 32, 33]. Due to the limited number of eligible studies, we only conducted meta-analyses on behalf of the presence of gallstones, as shown in Fig. 3C. The result still showed a significant association between the presence of gallstones and the risk of AOVC (OR=3.28; 95% CI 1.33-8.11; I^2=95.0%; p<0.001). In the subgroup analyses, the magnitudes of association were significantly higher in the Asian studies (OR=7.23; 95% CI 2.49-21.00; I^2=88.0%; p=0.004) than in the non-Asian studies (OR=1.57; 95% CI 1.28-1.92; I^2=0.0%; p=0.608) and in the studies that measured gallstones by imaging modality (OR=7.23; 95% CI 2.49-21.00; I^2=88.0%; p=0.004) than the studies that did not (OR=1.57; 95% CI 1.28-1.92; I^2=0.0%; p=0.608) (Table 5).

3.6. Sensitivity analysis and publication bias

The sensitivity analyses on the relationships between the presence of gallstones and the risk of BTC are given as Supplementary Fig. 1. We found similar results to the original meta-analysis result in the same directions and magnitudes of effects (SE ranging from 3.96 to 4.72 and each SE with the 95% CI embodying the original SE, 4.38) when we sequentially excluded every study one by one. The funnel plots referring to the association between the presence of gallstones and the risk of BTC by each subsite revealed no evidence for publication bias (Supplementary Fig. 2.). Egger’s tests did not identify a publication bias in the overall meta-analysis including all the subsites of
BTC (BTC, $t=0.79$, $p=0.421$; GBC, $t=1.98$, $p=0.068$; EBDC, $t=0.41$, $p=0.688$; AOVC: $t=1.13$, $p=0.340$) (Supplementary Fig. 2A-D.).
4. Discussion

Our systematic review and meta-analysis provided the most comprehensive evidence to date on the associations between gallstones and the risk of BTC, including GBC, EBDC, and AOVC. This study showed that the risks of GBC, EBDC, and AOVC increased with gallstone presence, and statistically significant associations were observed in both seven cohort studies and 19 case-control studies. In terms of gallstone size and number, the meta-analyses revealed that only size (>1 cm vs. <1 cm) was significantly associated with the risk of GBC. Sensitivity analyses of the restricted studies according to the study quality or adjustments as well as sequentially excluding any one study supported the stability of the results in the study.

Based on the meta-analysis results of the BTC subsites, specifically GBC and AOVC, the common trend of significantly stronger summary effect sizes on the association between the presence of gallstones and the risk of cancer was present among Asian studies and the studies that measured gallstones with various imaging modalities (ultrasonography, computed tomography, magnetic resonance imaging, and endoscopic retrograde cholangiopancreatography) than among their counterpart groups (Table 3, 5). This phenomenon should be further researched to understand the reason for the difference between the regional sectors, and it emphasizes the importance of identifying solid evidence of gallstones in the context of a preventive approach to BTC. In addition, the summary risk estimate of GBC in accordance with the gallstones' presence, although statistically insignificant, was the strongest among the subsites of BTC (Table 3-5), which resonated with the settled consensus [5]. The gallbladder carrying larger gallstones or crammed with multiple gallstones, already known to increase the risk of GBC [10], was also verified through the findings of our meta-analyses (S1 Table).
The previous reports that were reviewed altogether indicated that having gallstones was associated with an increased risk of BTC [32, 34, 35] and each subsite of BTC: GBC [27, 28, 30, 32-43], EBDC [12, 31-34, 39, 40, 44-47], and AOVC [13, 27, 30, 32, 33], although some studies reported a nonsignificant association [48, 49]. Our study summarized the results of these studies to obtain consistent results. However, the definition of the presence of gallstones deferred throughout the studies because the criteria were obscured or varied in the matter of deciding the minimal required lengths of time intervals between the gallstone-having status and the diagnosis of BTC. Some studies examined the presence of gallstones up to one year before the cancer diagnosis [30, 31], while others examined the presence of gallstones up to three years [32, 44] or more than one year [47] before the cancer diagnosis. Even in the two studies [35, 42], a lifetime history of gallstones was regarded as the presence of gallstones. Similarly, there is a paucity of studies that reported the duration of the presence of gallstones [27], which confines further implications on the relationships between the presence of gallstones and the carcinogenetic processes of BTC. The obscurity in defining the presence of gallstones and the lack of additional information on the attributes of gallstones, such as duration, may have contributed to the high heterogeneity within our meta-analyses.

With respect to the high heterogeneity of the included studies in our meta-analyses, no single factor among the study design, sex, geographic area, study period, measurement of exposure, study quality, and adjustments of confounders dramatically reduced the heterogeneity when applying the subgroup analysis. A notable finding is that the cohort study design, male sex, and measurement of gallstones with imaging study slightly alleviated the heterogeneity in the main analysis (Table 2). Similar trends were observed in the subsite analyses (Table 3, 4). This finding implies that the cohort studies adopting relatively objective methods for gallstone measurement reported much more precise and stable effect
sizes. For the stratification by sex where the degree of heterogeneity decreased across the male-grouped studies, a possible explanation may root in the unique epidemiological nature of cholelithiasis and BTC that the female sex and its related attributes (sex hormones, parity, and the number of pregnancies) are imperative risk factors for both diseases [10]. Contrary to the male groups, the female groups are under additional potential confounders, which are mostly unadjusted in previous studies. Thus, determining the association between the presence of gallstones and the risk of BTC has become much more complex, and the effect size of each study may tend to vary substantially.

The biological mechanisms linking gallstones and the risk of BTC are not well known. One hypothesis suggests that gallstones dropped down from the upstream biliary tract might result in chronic inflammation of the bile duct epithelium as underlying conditions for tumor development. That is, gallstones could lead to EBDC, causing inflammation of the bile duct wall [11]. In addition, approximately 35% of patients with stones develop complications such as cholecystitis or cholangitis [50], which may contribute to carcinogenesis in the gallbladder or bile ducts. Another possible hypothesis of the pathogenesis assumes that hormonal or reproductive factors might play a role in tumor development [51]. The increased exposure to endogenous estrogen and progesterone during pregnancy or exogenous estrogen seems to promote the formation of biliary stones. Under hormonal exposure, cholesterol saturation of bile mounts and that leads to impaired contractility of the smooth muscles of the biliary tract [52]. Therefore, biliary stasis and gallstone formation easily occur, which might be the key steps in the process of carcinogenesis in the biliary tract. [52].

Our meta-analysis results are not contradictory to either of these hypotheses.

There are several limitations to this systematic review and meta-analysis. First, we tried to capture the association between gallstones and the risk of BTC, thereby inevitably excluding some other studies [53-56] that provided the association between gallbladder...
disease (or condition), not gallstones, and the risk of BTC. Second, the definition of EBDC used in our study encompassed not only the EBDC and its equivalent term, EHC but also the CCA and its equivalent term, BDC. CCA (or BDC) is an overlapping term with EBDC, approximately 90% of which is EBDC, but the remaining 8 to 10% is intrahepatic bile duct cancer (IBDC), usually not a subsite of BTC [15, 57]. Third, although we extracted the risk estimates considering adjustments for the potential confounders, the scope of adjusted confounders varied across the studies, which could have caused a deviation of the meta-analysis results. Finally, there were significant heterogeneities across the studies, which would cast some doubts on the reliability of the summary risk estimates. These high heterogeneities may originate from the obscurity in defining gallstones' presence in previous studies so far, where most studies lacked concrete information about the duration of gallstones. These imply that the interval between the presence of gallstones and diagnosis of BTC is inconsistent among studies, leaving the same limitations for the meta-analyses. Therefore, future research needs to elicit clear criteria for gallstone presence, assuming that differences in the definition are a plausible source of the heterogeneity. Another reason for this phenomenon is that our meta-analyses combined all the eligible studies, in fact, have distinct natures. In our subgroup meta-analyses, the groups which share common study design (cohort study), sex (male), and measurement of exposure (imaging study) resulted in a reduced heterogeneity respectively compared to each of their counterparts. Further study with a more sophisticated approach is needed to narrow these specific groups to secure low en level of heterogeneity when synthesizing the risk estimates on the association between gallstones and BTC.

Despite these limitations, our study has several strengths. To the best of our knowledge, this study is the first systematic review and meta-analysis on the associations between gallstone characteristics and the risk of BTC. Unlike the previous systematic review [14],
we reported the characteristics of gallstones (presence, size, and number), not gallbladder disease as a whole, in association with the risk of BTC. Moreover, we conducted meta-analyses stratified by each subsite of BTC (GBC, EBDC, and AOVC) and other diverse factors, including the study design, sex, geographic area, study period, measurement of exposure, study quality, and whether analyses were adjusted for various confounders. In this study, we attempted to explore all the relevant studies and reflect the findings and achievements hitherto established to the greatest extent capable.

In conclusion, we found statistically significant associations between gallstones and the increased risk of BTC through systematic reviews and meta-analyses. We verified that the presence of gallstones is a critical risk factor for BTC as well as for GBC, EBDC, and AOVC. Our study provides a better description of the association between gallstones and the risk of BTC.

Ethics Statement
Informed consent was waived due to the study design: systematic review and meta-analysis.

Conflicts of Interest
The authors have no conflicts of interest to declare for this study

Acknowledgements
This work was supported by a grant from the Seoul National University Hospital (2018, 2020).

Author Contributions
Conceptualization: DH, HJ, SC, NS, AS. Data curation: DH, HJ, NS, WK. Formal analysis: HJ, DH, NS, SC, AS. Funding acquisition: AS. Methodology: DH, HJ, NS, SC. Project administration: AS. Visualization: DH, HJ. AS. Writing – original draft: HJ, DH. Writing – review & editing: HJ, DH, NS, SC, AS.

Reference


Table 1. Characteristics of studies included in the meta-analysis

<table>
<thead>
<tr>
<th>No.</th>
<th>First author (year)</th>
<th>Country</th>
<th>Study period</th>
<th>Sample size</th>
<th>Outcome</th>
<th>Gallstone characteristics</th>
<th>Gallstone Measurement</th>
<th>Matching</th>
<th>Adjustment</th>
<th>NOS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Chow (1999)</td>
<td>Denmark</td>
<td>1977-1993</td>
<td>17715 (42)</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(12)</td>
<td>EBDC</td>
<td>Present vs. absent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(8)</td>
<td>AOVVC</td>
<td>Present vs. absent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Zou (2000)</td>
<td>China</td>
<td>1986-1998</td>
<td>105019 (3922)</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(93)</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(142)</td>
<td>EBDC</td>
<td>Present vs. absent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Nordenstedt (2012)</td>
<td>Sweden</td>
<td>1965-2008</td>
<td>192960 (169)</td>
<td>EBDC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex, calendar year</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Lai (2013)</td>
<td>Taiwan</td>
<td>1996-2008</td>
<td>DM</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex</td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Case-control study</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8 Diehl (1983) U.S.</td>
<td>1976-1980</td>
<td>45/66 GBC</td>
<td>Size (≥1 vs. &lt;1)</td>
<td>Medical record</td>
<td>Age, sex, hospital, screening year</td>
<td></td>
</tr>
<tr>
<td>9 Lowenfels (1985) U.S.</td>
<td>Black 1965-1978</td>
<td>74/2013 GBC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td></td>
<td>White 1977-1983</td>
<td>57/386 GBC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Lowenfels (1989) U.S.</td>
<td>1979-1985</td>
<td>15/398 GBC</td>
<td>Size (≥1 vs. &lt;1)</td>
<td>Medical record</td>
<td>Age, race</td>
<td></td>
</tr>
<tr>
<td>11 Moerman (1993) Netherlands</td>
<td>1966-1969</td>
<td>43/98 GBC</td>
<td>Size (≥1 vs. &lt;1)</td>
<td>Medical record</td>
<td>Age, sex, hospital, date of admission</td>
<td></td>
</tr>
<tr>
<td></td>
<td>41/96 GBC</td>
<td>Size (≥2 vs. &lt;2)</td>
<td>Medical record</td>
<td>Age, sex, hospital, with imaging date of admission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Years</td>
<td>n</td>
<td>Cause</td>
<td>Outcome</td>
<td>Controls</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>---</td>
<td>-------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>14 Welzel (2007)</td>
<td>U.S.</td>
<td>1993-1999</td>
<td>549/102782</td>
<td>EHC</td>
<td>Present vs. absent</td>
<td>Medical record</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BDC</td>
<td>Present vs. absent</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AOVC</td>
<td>Present vs. absent</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>No.</th>
<th>Study (Year)</th>
<th>Country</th>
<th>Period</th>
<th>Cases/Total</th>
<th>Disease</th>
<th>Recurrence</th>
<th>Follow-up</th>
<th>Variables</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Grainge (2009)</td>
<td>U.K.</td>
<td>1987-2002</td>
<td>372/5760 CCA 184/5760 GBC</td>
<td>Present vs. absent Present vs. absent</td>
<td>Medical record Sex, age, GP practice</td>
<td>Alcohol, smoking, BMI</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Tao (2010)</td>
<td>China</td>
<td>1998-2008</td>
<td>129/380 EHC</td>
<td>Present vs. absent</td>
<td>Medical record with imaging</td>
<td>Age, sex Age, sex, DM, history of cholecystectomy</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Alvi (2011)</td>
<td>Pakistan</td>
<td>1988-2007</td>
<td>60/120 GBC</td>
<td>Size (≥1 vs. &lt;1) Number (&gt;1 vs. 1)</td>
<td>Medical record with imaging</td>
<td>Age, sex Age, parity, BMI, stone characteristics</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Wu (2012)</td>
<td>China</td>
<td>1998-2010</td>
<td>93/809 GBC 86/835 EBDC</td>
<td>Present vs. absent Present vs. absent</td>
<td>Medical record with imaging</td>
<td>Age, sex Age, sex, HBV, DM, TC, HDL-C</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Onal (2012)</td>
<td>Turkey</td>
<td>2006-2010</td>
<td>99/48 CCA</td>
<td>Present vs. absent</td>
<td>Self-reported</td>
<td>Age, sex</td>
<td>Age, sex, HBV, alcohol, smoking</td>
<td>6</td>
</tr>
<tr>
<td>23</td>
<td>Chang (2013)</td>
<td>Taiwan</td>
<td>2004-2008</td>
<td>2179/8716 EHC</td>
<td>Present vs. absent</td>
<td>Medical record with imaging</td>
<td>Age, sex, date of diagnosis Cholangitis</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Nogueira (2014)</td>
<td>U.S.</td>
<td>1992-2005</td>
<td>3681/100000 (3664) GBC (1646) AOV</td>
<td>Present vs. absent Present vs. absent Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex, calendar year</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>He (2014)</td>
<td>China</td>
<td>2006-2010</td>
<td>210/62 AOV</td>
<td>Present vs. absent</td>
<td>Medical record</td>
<td>Age, sex</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>No.</th>
<th>Study Year</th>
<th>Country</th>
<th>Study Period</th>
<th>No. of Participants</th>
<th>Cancer Site</th>
<th>Present vs. Absent</th>
<th>Medical Record w/ Imaging</th>
<th>Age, Sex</th>
<th>Hypertension, DM, VOD, Alcohol, Smoking, BMI, PLG</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>Cha (2015)</td>
<td>Korea</td>
<td>2008-2013</td>
<td>78/78</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Medical record with imaging</td>
<td>Age, sex</td>
<td>Hypertension, DM, VOD, alcohol, smoking, BMI, PLG</td>
</tr>
<tr>
<td>27</td>
<td>Lee (2015)</td>
<td>Korea</td>
<td>2007-2013</td>
<td>276/552</td>
<td>CCA</td>
<td>Present vs. absent</td>
<td>Medical record w/ Imaging</td>
<td>Age, sex, date of diagnosis</td>
<td>Alcohol, DM, HBV, LFI</td>
</tr>
<tr>
<td>28</td>
<td>Lee (2015)</td>
<td>Korea</td>
<td>2007-2013</td>
<td>81/162</td>
<td>EHC</td>
<td>Present vs. absent</td>
<td>Medical record w/ Imaging</td>
<td>Age, sex, date of diagnosis</td>
<td>DM, smoking</td>
</tr>
<tr>
<td>29</td>
<td>Rosato (2016)</td>
<td>Italy</td>
<td>Study 1 (1983-1992)</td>
<td>159/795</td>
<td>BTC</td>
<td>Present vs. absent</td>
<td>Self-reported</td>
<td>Age, sex, study, center</td>
<td>Year of interview, education, BMI, alcohol, smoking</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Study 2 (1994-2009)</td>
<td></td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Self-reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Tamrakar (2016)</td>
<td>Nepal</td>
<td>2012-2013</td>
<td>100/100</td>
<td>GBC</td>
<td>Present vs. absent</td>
<td>Self-reported</td>
<td>Age, sex, marital status</td>
<td>Education, hospital, smoking, fruit consumption, residence</td>
</tr>
</tbody>
</table>

NOS, newcastle-ottawa scale; AOVC, ampulla of vater cancer; BDC, bile duct cancer; BMI, body mass index; BTC, biliary tract cancer; CCA, cholangiocarcinoma; DM, diabetes mellitus; EBDC, extrahepatic bile duct cancer; EHC, extrahepatic cholangiocarcinoma; GBC, gallbladder cancer; GP, general physician; HBV, hepatitis b virus; HDL-C, high density lipoprotein cholesterol; LFI, liver fluke infestation; PLG, polypoid lesion of gallbladder; VOD, vascular occlusive disease.

1 Number of participants (case) for cohort study and case/control (case of sub-site) for case-control study. 2 NOS for assessing the quality of non-randomised studies in meta-analysis. 3
Gallstone presence of up to 3 years before diagnosis of cancer. Gallstone presence more than 1 year before diagnosis of cancer. Gallstone presence of up to 1 year before diagnosis of cancer. History of gallstone (ever vs. never)
Table 2. Meta-analysis results for association between the presence of gallstone and the risk of BTC by the subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. of study</th>
<th>OR (95%CI) (^1)</th>
<th>(I^2) value (%)</th>
<th>(P) for heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td>26</td>
<td>4.38 (3.23-5.93)</td>
<td>91.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study</td>
<td>7</td>
<td>3.17 (2.28-4.39)</td>
<td>79.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Case-control study</td>
<td>19</td>
<td>5.04 (3.36-7.56)</td>
<td>90.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>9</td>
<td>3.40 (2.70-4.28)</td>
<td>35.8</td>
<td>0.132</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>4.26 (2.75-6.59)</td>
<td>84.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Geographic area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>15</td>
<td>5.25 (3.50-7.86)</td>
<td>82.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-Asia (^2)</td>
<td>11</td>
<td>3.58 (2.17-5.91)</td>
<td>95.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Study period (^3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 2000</td>
<td>8</td>
<td>5.39 (2.57-11.34)</td>
<td>95.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Around 2000</td>
<td>7</td>
<td>2.67 (2.10-3.39)</td>
<td>38.6</td>
<td>0.135</td>
</tr>
<tr>
<td>After 2000</td>
<td>7</td>
<td>5.21 (2.13-12.74)</td>
<td>85.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No record</td>
<td>4</td>
<td>5.73 (2.61-12.61)</td>
<td>87.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Measure of gallstone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical record with imaging study</td>
<td>8</td>
<td>7.09 (3.87-12.98)</td>
<td>64.5</td>
<td>0.004</td>
</tr>
<tr>
<td>Medical record without imaging study</td>
<td>16</td>
<td>3.81 (2.48-5.85)</td>
<td>93.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No record</td>
<td>2</td>
<td>3.47 (2.88-4.18)</td>
<td>17.1</td>
<td>0.272</td>
</tr>
<tr>
<td>Study quality (^4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High NOS</td>
<td>14</td>
<td>3.99 (2.85-5.59)</td>
<td>75.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low NOS</td>
<td>12</td>
<td>4.81 (2.87-8.05)</td>
<td>94.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for age, yes</td>
<td>21</td>
<td>3.71 (2.66-5.16)</td>
<td>90.6</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\(^1\) OR: Odds Ratio; CI: Confidence Interval

\(^2\) Asia: China, Japan, Korea, Thailand, India, Singapore, Malaysia, Indonesia, the Philippines, and Vietnam; Non-Asia: the rest of the world


\(^4\) High NOS: a score of 7 or above; Low NOS: a score of 6 or below

(Continued)
<table>
<thead>
<tr>
<th>Adjustment for sex, yes</th>
<th>20</th>
<th>3.92 (2.77-5.55)</th>
<th>91.2</th>
<th>&lt;0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjustment for comorbidities, yes</td>
<td>7</td>
<td>3.05 (1.84-5.05)</td>
<td>75.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for lifestyle factors, yes</td>
<td>4</td>
<td>4.84 (1.95-11.98)</td>
<td>85.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for education, yes</td>
<td>1</td>
<td>9.42 (3.56-24.91)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Adjustment for geographic areas, yes</td>
<td>4</td>
<td>7.34 (2.28-23.62)</td>
<td>87.6</td>
<td>0.000</td>
</tr>
</tbody>
</table>

BTC, biliary tract cancer; OR, odds ratio; NOS, Newcastle-Ottawa Scale.

1 OR (Odds ratio) refers to summary estimate of effects based on random effects model. 2 Non-Asia including U.S. and European areas. 3 Study period is defined by the study’s starting point (a) and ending point (b). Before 2000, (a) and (b) are both before 2000; Around 2000, (a) is before 2000 but (b) is after 2000; After 2000, (a) and (b) are both after 2000. Non-Asia including U.S. and European areas. 4 The quality scores equal or more than the median value was judged as high NOS (≥7). 4 Adjustment for lifestyle factors such as alcohol, smoking, BMI, or etc.
Table 3. Meta-analysis results for association between the presence of gallstone and the risk of GBC by the subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. of study</th>
<th>OR (95%CI)</th>
<th>I² value (%)</th>
<th>P for heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td>16</td>
<td>7.26 (4.33-12.18)</td>
<td>93.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study</td>
<td>5</td>
<td>4.54 (2.62-7.87)</td>
<td>72.5</td>
<td>0.006</td>
</tr>
<tr>
<td>Case-control study</td>
<td>11</td>
<td>9.60 (4.45-20.70)</td>
<td>95.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>6.69 (3.13-14.31)</td>
<td>74.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>6.69 (2.46-18.15)</td>
<td>87.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Geographic area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>8</td>
<td>12.72 (6.35-25.46)</td>
<td>86.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-Asia¹</td>
<td>8</td>
<td>3.59 (2.68-4.81)</td>
<td>56.0</td>
<td>0.026</td>
</tr>
<tr>
<td>Study period</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 2000</td>
<td>6</td>
<td>6.22 (3.24-11.95)</td>
<td>57.3</td>
<td>0.039</td>
</tr>
<tr>
<td>Around 2000</td>
<td>5</td>
<td>4.54 (1.75-11.81)</td>
<td>97.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>After 2000</td>
<td>2</td>
<td>33.59 (7.96-141.72)</td>
<td>0.0</td>
<td>0.611</td>
</tr>
<tr>
<td>No record</td>
<td>3</td>
<td>11.66 (2.70-50.45)</td>
<td>96.3</td>
<td>0.000</td>
</tr>
<tr>
<td>Measure of gallstone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical record with imaging</td>
<td>5</td>
<td>15.27 (7.48-31.18)</td>
<td>76.9</td>
<td>0.002</td>
</tr>
<tr>
<td>Medical record without imaging</td>
<td>11</td>
<td>4.67 (3.29-6.61)</td>
<td>76.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Study quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High NOS</td>
<td>8</td>
<td>6.72 (2.80-16.10)</td>
<td>94.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low NOS</td>
<td>8</td>
<td>7.25 (4.08-12.88)</td>
<td>83.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for age, yes</td>
<td>12</td>
<td>6.90 (3.80-12.51)</td>
<td>94.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for sex, yes</td>
<td>11</td>
<td>7.83 (4.11-14.91)</td>
<td>95.2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

(Continued)
| Adjustment for comorbidities, yes | 3 | 3.99 (2.14-7.45) | 27.5 | 0.252 |
| Adjustment for lifestyle factors, yes | 6 | 6.87 (3.36-14.04) | 69.9 | 0.005 |
| Adjustment for education, yes | 1 | 23.80 (17.00-33.32) | - | - |
| Adjustment for geographic areas, yes | 4 | 9.32 (2.95-29.37) | 76.7 | 0.005 |

GBC, gallbladder cancer; OR, odds ratio; NOS, Newcastle-Ottawa Scale.

1 OR (Odds ratio) refers to summary estimate of effects based on random effects model. 2 Non-Asia including U.S. and European areas. 3 Study period is defined by the study’s starting point (a) and ending point (b). Before 2000, (a) and (b) are both before 2000; Around 2000, (a) is before 2000 but (b) is after 2000; After 2000, (a) and (b) are both after 2000. 4 The quality scores equal or more than the median value was judged as high NOS (≥7). 5 Adjustment for lifestyle factors such as alcohol, smoking, BMI, or etc.
Table 4. Meta-analysis results for association between presence of gallstone and the risk of EBDC by subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. of study</th>
<th>OR (95%CI) ¹</th>
<th>I² value (%)</th>
<th>P for heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td>17</td>
<td>3.17 (2.24-4.50)</td>
<td>95.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Subsite</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBDC (or EHC)</td>
<td>12 ²</td>
<td>2.87 (2.06-3.99)</td>
<td>95.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CCA (or BDC)</td>
<td>6 ²</td>
<td>2.12 (1.35-3.33)</td>
<td>92.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Study design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study</td>
<td>4</td>
<td>2.33 (2.00-2.72)</td>
<td>21.4</td>
<td>0.282</td>
</tr>
<tr>
<td>Case-control study</td>
<td>13</td>
<td>3.67 (2.26-5.95)</td>
<td>96.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
<td>3.46 (2.29-5.22)</td>
<td>78.6</td>
<td>0.001</td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>5.13 (2.73-9.66)</td>
<td>91.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Geographic area</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>9</td>
<td>3.48 (2.30-5.28)</td>
<td>83.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-Asia</td>
<td>8</td>
<td>2.99 (1.72-5.19)</td>
<td>97.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Study period</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 2000</td>
<td>5</td>
<td>4.43 (1.61-12.20)</td>
<td>97.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Around 2000</td>
<td>5</td>
<td>2.79 (1.80-4.32)</td>
<td>93.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>After 2000</td>
<td>4</td>
<td>2.28 (0.79-6.55)</td>
<td>82.0</td>
<td>0.001</td>
</tr>
<tr>
<td>No record</td>
<td>3</td>
<td>3.50 (1.66-7.35)</td>
<td>88.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Measure of gallstone</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical record with imaging</td>
<td>4</td>
<td>5.52 (3.07-9.93)</td>
<td>69.1</td>
<td>0.021</td>
</tr>
<tr>
<td>study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical record without imaging</td>
<td>11</td>
<td>2.60 (1.61-4.20)</td>
<td>96.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No record</td>
<td>2</td>
<td>3.47 (2.88-4.18)</td>
<td>17.1</td>
<td>0.272</td>
</tr>
<tr>
<td><strong>Study quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High NOS</td>
<td>9</td>
<td>3.64 (2.37-5.59)</td>
<td>84.9</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Low NOS</th>
<th>8</th>
<th>2.79 (1.58-4.93)</th>
<th>97.5</th>
<th>&lt;0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjustment for age, yes</td>
<td>16</td>
<td>3.26 (2.23-4.78)</td>
<td>95.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for sex, yes</td>
<td>15</td>
<td>3.34 (2.26-4.94)</td>
<td>95.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for comorbidities, yes</td>
<td>5</td>
<td>2.60 (1.39-4.89)</td>
<td>85.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adjustment for lifestyle factors, yes</td>
<td>4</td>
<td>2.12 (0.83-5.44)</td>
<td>82.9</td>
<td>0.001</td>
</tr>
<tr>
<td>Adjustment for education, yes</td>
<td>1</td>
<td>8.00 (5.60-11.43)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Adjustment for geographic areas, yes</td>
<td>3</td>
<td>3.92 (1.16-13.29)</td>
<td>75.0</td>
<td>0.018</td>
</tr>
</tbody>
</table>

BDC, bile duct cancer; CCA, cholangiocarcinoma; EBDC, extrahepatic bile duct cancer; EHC, extrahepatic cholangiocarcinoma; OR, odds ratio; NOS, Newcastle-Ottawa Scale.

1 OR (Odds ratio) refers to summary estimate of effects based on random effects model. 2 One study [30] suggested the risk estimates of both EBDC and CCA. 3 Non-Asia including U.S. and European areas. 4 Study period is defined by the study’s starting point (a) and ending point (b). Before 2000, (a) and (b) are both before 2000; Around 2000, (a) is before 2000 but (b) is after 2000; After 2000, (a) and (b) are both after 2000. 5 The quality score equal or more than median value was judged as high NOS (≥7). 6 Adjustment for lifestyle factors such as alcohol, smoking, BMI, or etc.
Table 5. Meta-analysis results for association between presence of gallstone and the risk of AOVC by subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>No. of study</th>
<th>OR (95%CI)</th>
<th>$I^2$ value (%)</th>
<th>$P$ for heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td>5</td>
<td>3.28 (1.33-8.11)</td>
<td>93.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study</td>
<td>1</td>
<td>2.30 (1.00-5.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case-control study</td>
<td>4</td>
<td>3.56 (1.20-10.54)</td>
<td>95.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1</td>
<td>3.60 (1.70-7.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>3.30 (1.41-7.70)</td>
<td>43.2</td>
<td>0.185</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>2</td>
<td>7.23 (2.49-21.00)</td>
<td>88.0</td>
<td>0.004</td>
</tr>
<tr>
<td>Non-Asia</td>
<td>3</td>
<td>1.57 (1.28-1.92)</td>
<td>0.0</td>
<td>0.608</td>
</tr>
<tr>
<td>Study period</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 2000</td>
<td>1</td>
<td>1.88 (0.61-5.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Around 2000</td>
<td>2</td>
<td>2.46 (0.91-6.61)</td>
<td>92.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>After 2000</td>
<td>1</td>
<td>12.47 (7.37-21.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No record</td>
<td>1</td>
<td>2.30 (1.00-5.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measure of gallstone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical record with imaging</td>
<td>2</td>
<td>7.23 (2.49-21.00)</td>
<td>88.0</td>
<td>0.004</td>
</tr>
<tr>
<td>Medical record without imaging</td>
<td>3</td>
<td>1.57 (1.28-1.92)</td>
<td>0.0</td>
<td>0.608</td>
</tr>
<tr>
<td>Study quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High NOS</td>
<td>4</td>
<td>4.18 (1.79-9.80)</td>
<td>83.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low NOS</td>
<td>1</td>
<td>1.52 (1.23-1.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustment for age, yes</td>
<td>4</td>
<td>2.29 (1.26-4.17)</td>
<td>77.0</td>
<td>0.005</td>
</tr>
<tr>
<td>Adjustment for sex, yes</td>
<td>3</td>
<td>2.39 (1.17-4.91)</td>
<td>84.6</td>
<td>0.001</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---</td>
<td>------------------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Adjustment for education, yes</td>
<td>1</td>
<td>4.20 (2.50-7.06)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

AOVC, ampulla of Vater cancer; OR, odds ratio; NOS, Newcastle-Ottawa Scale.

1 OR (Odds ratio) refers to summary estimate of effects based on random effects model. 2 Non-Asia including U.S. and European areas. 3 Study period is defined by the study’s starting point (a) and ending point (b). Before 2000, (a) and (b) are both before 2000; Around 2000, (a) is before 2000 but (b) is after 2000; After 2000, (a) and (b) are both after 2000. 4 The quality score equal or more than median value was judged as high NOS (≥7).
Fig. 1. Flow chart of study selection for the meta-analysis.
Fig. 2. Forest plot showing the relationship between presence of gallstone and the risk of BTC

1 Effect size of Lai et al. was calculated by pooling the results of DM group and non-DM group.
2 Effect size of Lowenfels et al. was calculated by pooling the results of Indian and non-Indian.
3 OR (Odds ratio) and 95% CI (Confidence interval) refers to the estimate of effects included in a random effects model.
** Pooled the results of two types BTC subsites.
*** Pooled the results of three types BTC subsites.
Fig. 3A. Forest plot showing the relationship between presence of gallstone and the risk of GBC

1 Effect size of Lai et al. was calculated by pooling the results of DM group and non-DM group.
2 Effect size of Lowenfels et al. was calculated by pooling the results of Indian and non-Indian group.
3 OR (Odds ratio) and 95% CI (Confidence interval) refers to the estimate of effects included in a random effects model.
Fig. 3B. Forest plot showing the relationship between presence of gallstone and the risk of EBDC

1 Pooled the risk estimates of cholecystolithiasis and choledocholithiasis group.
2 Pooled the risk estimates of CCA and EBDC group.
3 OR (Odds ratio) and 95% CI (Confidence interval) refers to the estimate of effects included in a random effects model.
Fig. 3C. Forest plot showing the relationship between presence of gallstone and the risk of AOVC.

1 OR (Odds ratio) and 95% CI (Confidence interval) refers to the estimate of effects included in a random effects model.