Clinical efficacy of adjuvant treatments for patients with resected biliary tract cancer: a systematic review and network meta-analysis

Ye Chen,1 Baoxia Zhang,2 Chang Liu,2 Ye Cao,3 Cheng Lyu,2 Meng Qiu1

ABSTRACT

Objective This study aimed to determine the benefits of adjuvant therapy in patients with resected biliary tract cancer (BTC) and identify the optimal adjuvant treatment scheme.

Design Systematic review and network meta-analysis.

Data sources Studies comparing different adjuvant therapies in patients with BTC were searched in PubMed, Embase, CINAHL, Cochrane Central Register of Controlled Trials and ClinicalTrials.gov databases from inception to December 2021.

Materials and methods Eligible studies were identified, and data were extracted independently by two authors. A random-effects network meta-analysis was performed using R software. The pooled outcomes of overall survival (OS) and disease-free survival (DFS) were measured using the combined HRs with 95% CIs.

Results Nineteen eligible studies reporting three types of adjuvant therapies were included in our network meta-analysis. Adjuvant radiotherapy (ART, HR 0.62; 95% CI 0.42 to 0.93), adjuvant chemoradiotherapy (ACRT; HR 0.71; 95% CI 0.54 to 0.83) and adjuvant chemotherapy (ACT; HR 0.84; 95% CI 0.68 to 0.98) showed improvement in OS, but not in DFS. Due to the lack of head-to-head studies of ART, ACRT and ACT, the above results need to be further verified by prospective randomised controlled trials.

Introduction Biliary tract cancer (BTC) is classified into gallbladder cancer (GBC), intrahepatic cholangiocarcinoma (iCCA) and extrahepatic cholangiocarcinoma (eCCA), and eCCA is further subdivided into perihilar cholangiocarcinoma (pCCA or Klatskin tumour) and distal cholangiocarcinoma (dCCA). BTC accounts for approximately 3% of digestive system cancers and 10%–15% of primary liver cancers.1 The incidence of BTC is higher in Asian countries and lower in European countries, the USA and Australia; however, its incidence is increasing globally.2,3 The vast majority (>90%) of BTC cases were adenocarcinoma. Moreover, most cases are usually in the advanced or metastatic stage at the initial diagnosis. Only approximately 20% of BTC cases are considered resectable.4 Surgical resection remains the primary curative treatment in patients with resected BTC.5–8
BTC. However, the high recurrence rates (including locoregional or distant recurrence) and low survival rates (5-year survival rates of patients ranging from 5% to 15%) are prominent problems, even with complete (R0) resection.4–8 Previously published studies have shown that histologic margin status, lymph node (LN) involvement and intrahepatic metastasis are the main prognostic factors.9–11 Adjuvant chemotherapy (ACT), adjuvant radiotherapy (ART) and adjuvant chemoradiotherapy (ACRT) are the main options following resection for adjuvant therapy.

A variety of guidelines have suggested that postoperative adjuvant therapy could be considered an option for BTC patients.12–15 The use of ACT and ACRT for patients with GBC and eCCA was supported by the National Comprehensive Cancer Network (NCCN) guidelines. Adjuvant capecitabine is recommended by the American Society of Clinical Oncology (ASCO) guidelines for patients with resected BTC for 6 months, based on the results of the BILCAP study, which showed a protocol-specified adjusted overall survival (OS) HR of 0.71 (95% 0.55 to 0.92).16 However, some experts still hold reservations regarding the BILCAP study due to the underpowered statistical design and concerns over data maturity.17 The discrepancy between the BILCAP study and the PRODIGE-12/ACCORD-18 studies also results in more discussions on the optimal scheme of adjuvant therapy.18 Despite the tremendous effort, universal agreement on the optimal scheme of adjuvant therapy has not been established due to the lack of high-quality evidence.19

Moreover, conflicting results were also observed in the meta-analyses on this topic. A previous study by Rangarajan et al showed a significant improvement in OS with adjuvant therapy after resection compared with that of resection only.20 In contrast, Horgan et al reported a non-significant improvement in OS with adjuvant therapy compared with that of observation in the overall population. However, a more significant benefit from adjuvant therapy was found in patients with positive LN involvement and R1 resection.21 Three network meta-analyses of adjuvant therapies in patients with resected BTC have been published previously. One made comparisons among ACT, ACRT and resection only and did not consider ART.22 One only summarised three randomised clinical trials (RCTs) covering three different ACT regimens, and the robustness of their results was limited due to the very small sample sizes.23 A recent study indicated that ACRT could provide a survival benefit in patients with positive margins or nodal involvement, but the inclusion of patients with R2 margins is still debatable.24

Therefore, it is essential to perform a new network meta-analysis to elucidate the efficacy of adjuvant therapy and identify the optimal scheme of adjuvant therapy for patients with resected BTC.

**METHODS**

The study design was built in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyse extension statement for network meta-analysis for healthcare (online supplemental table 1).25

**Search strategy**

PubMed, Embase, CINAHL, Cochrane Central Register of Controlled Trials (CENTRAL) and ClinicalTrials.gov databases were searched from inception to December 2021 to find relevant literature using the main keywords “biliary tract cancer” and “adjuvant therapy.” In addition, references in the relevant literature were manually checked for potentially relevant papers. The detailed search strategy is presented in online supplemental table 2.

**Selection criteria**

Studies that met the following criteria were included: (1) those including patients with histologically or cytologically confirmed resected BTC and treated with adjuvant therapy or observation after curative-intent resection (defined as negative resection margins (RM, R0) or microscopic positive resection margins (R1), but not macroscopic involvement resections (R2)); (2) studies reporting at least one of the following clinical outcomes, OS or disease-free survival (DFS) and (3) studies with treatment and control arms.

Studies that involved the following were excluded: (1) studies including patients with ampullary carcinomas or other primary cancers or neoadjuvant therapy, palliative therapy or therapy after postoperative recurrence; (2) studies comparing the same type of adjuvant therapies; (3) those with unbalanced baseline profiles in age, sex, disease severity (American Joint Committee on Cancer staging) or residual tumour status; (4) reviews, conference abstracts, posters or case reports and (5) studies not written in English or without a full text, or with a small sample size (<10 in any group).

All study titles and abstracts were screened, and then the full texts of potentially eligible articles were sequentially assessed for final inclusion.

**Data extraction and quality assessment**

The following details were extracted from each study: author, study period, country, study type, sample size, tumour site (BTC type), disease severity, age, female sex (%), RM status, LN status, OS and DFS. OS was defined as the period from the date of surgery (or randomisation in RCTs) to the date of death (or last follow-up). DFS was defined as the time from the surgery date (or randomisation in RCTs) to recurrence of tumours (locoregional or distant). Moreover, the authors were contacted when there was confusion or missing data in any article. Data were excluded or not considered if no response was received.

The quality of the studies was assessed using the Newcastle-Ottawa Scale for observational studies, based
on the following domains: selection, comparability and outcome. The detailed information can be found in online supplemental table 3. A study that scored 8–9 points, of which 2 points for comparability, is considered high quality; a score of 4–7 points indicates moderate quality and 0–3 means low quality. The Cochrane Risk of Bias Tool was used to grade the RCT quality.

Two investigators (YCh and BZ) independently conducted the study selection and data extraction, and four authors (YCh, BZ, CLi and YCa) assessed the risk of bias of each eligible study. Discrepancies were resolved by consensus and arbitration by other investigators (CLu and MQ).

**Data synthesis and statistical analysis**

The primary outcomes were OS and DFS and measured using the HR with a 95% CI. When HRs and 95% CIs for OS and DFS were not reported in the original article, they were extracted from survival curves using Engauge Digitizer V.10.9 (2014 Mark Mitchell) and estimated by the method suggested by Tierney. A Bayesian network meta-analysis was performed using the R software (V.4.1.2, https://www.r-project.org/). Network plots were generated for different outcomes to clarify which treatments were compared directly or indirectly. A random-effects model (including subgroup analyses based on the RM status, tumour sites, regions and patients without distal metastasis) was used to compare all direct and indirect evidence using ‘Rjags’ and ‘gemtc’ packages in the R software. To fit the non-informative uniform and normal prior distributions, the parameters were set with four chains, 50 000 sample iterations (n. inter), and 20 000 burn-ins (n.adapt) with a thinning interval of 1. The convergence of the chains was assessed using the Gelman–Rubin statistics and inspection of the trace plots (online supplemental figure 1). The deviance information criterion (DIC) was used to test the goodness of fit of consistent and inconsistent models.

The transitivity assumption was evaluated by comparing the distribution of clinical variables (age, percentage of females, sample size, publication year, RM status, LN status), which could be effect modifiers. The local inconsistency of the model was evaluated using the node-splitting approach. The global and local heterogeneity was assessed via between-study variance $\tau^2$ and $I^2$ inconsistency statistic. Heterogeneity was considered low, moderate, and high for estimated $\tau^2$ or $I^2$ values <25%, between 25% and 50%, and larger than 50%, respectively. Within the Bayesian framework, the network meta-analysis provided a ranking probability of each treatment and estimated the overall rankings by calculating the surface under the cumulative ranking (SUCRA). To assess the reliability of results, sensitivity analyses for OS and DFS were performed by excluding RCT studies, only including high-quality studies, and only including the studies for which HRs were reported in the original articles. The ‘netmeta’ package in the R software was used to generate the comparison-adjusted funnel plots to visualise publication bias. Egger’s test was used to assess funnel plot asymmetry. Values of p<0.05 were considered statistically significant. All statistical tests were two sided.

**Patient and public involvement**

Patients or the public were not involved in our research design, conduct, reporting or dissemination plans.

**RESULTS**

**Characteristics of the included studies**

A total of 2146 records were identified, including 49 records that were manually searched by reviewing the references of relevant publications. After reviewing the abstracts, 264 full-text articles were retrieved and reviewed after excluding 688 duplicates and 1194 ineligible records (figure 1). Nineteen studies comprising a total of 5595 patients who received one of the following four different treatments after radical resection—ACT, ART, ACRT and surgical resection alone (observation group)—were included in our analysis. It is worth noting that 34 studies were excluded because of unbalanced or unclear baselines. Of the included patients, 2774 patients received adjuvant treatments following curative-intent resection, and 2821 patients underwent curative-intent resection without adjuvant treatments.

The main characteristics of the studies included in the network meta-analysis are presented in table 1. These 19 studies consisted of 3 RCTs and 16 retrospective studies. The patients of 6 studies were Westerners (Americans, French and British) and 13 studies were Asian (Japanese, Korean, Indian and Chinese). Four studies assessed patients...
Table 1  Characteristics of included studies

<table>
<thead>
<tr>
<th>Studies</th>
<th>Country</th>
<th>Study period</th>
<th>BTC type</th>
<th>Disease severity</th>
<th>Therapy</th>
<th>Regimen of CT/ type of RT</th>
<th>Sample size</th>
<th>Margin status (R0/R1)</th>
<th>Nodal status (negative/positive)</th>
<th>Female no (%)</th>
<th>Age, mean (SD/range), years</th>
<th>5-year OS (%); MOST (months)</th>
<th>5-year DFS (%); MDST (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACT vs observation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kobayashi 2012</td>
<td>Japan</td>
<td>1989–2010</td>
<td>CCA</td>
<td>I–IV</td>
<td>ACT</td>
<td>GEM/S–1</td>
<td>51</td>
<td>N0:26; N+: 25</td>
<td>28 (55)</td>
<td>–</td>
<td>46.0; –</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Morine 2017</td>
<td>Japan</td>
<td>1995–2012</td>
<td>CCA/GBC</td>
<td>II–IVb (T1–T4)</td>
<td>ACT</td>
<td>GEM +5–FU+Cis</td>
<td>28</td>
<td>N0:31; N+: 23</td>
<td>14 (26)</td>
<td>–</td>
<td>23.0; –</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Akahoshi 2018</td>
<td>Japan</td>
<td>2004–2015</td>
<td>CCA</td>
<td>0–IVa (T1–T4)</td>
<td>ACT</td>
<td>GEM</td>
<td>26</td>
<td>N0:15/11</td>
<td>9 (35)</td>
<td>–</td>
<td>8.0; –</td>
<td>–</td>
<td>25.0; –</td>
</tr>
<tr>
<td>Bergeat 2018</td>
<td>France</td>
<td>2000–2015</td>
<td>dCCA</td>
<td>Ia–III</td>
<td>ACT</td>
<td>GEM</td>
<td>49</td>
<td>N0:12/37</td>
<td>7/42 (45)</td>
<td>18 (37)</td>
<td>62.0 (36–77); 70.9; 44.2;</td>
<td>25.6; –</td>
<td>15.5; 43.3</td>
</tr>
<tr>
<td>Miyata 2021</td>
<td>Japan</td>
<td>2007–2018</td>
<td>GBC/pCCA/</td>
<td>I–IV</td>
<td>ACT</td>
<td>S–1</td>
<td>38</td>
<td>N0:21; N1: 17</td>
<td>13 (34)</td>
<td>72 (52–82); 71.0; 51.6; 51.6;</td>
<td>–</td>
<td>15.6; 61.2</td>
<td></td>
</tr>
<tr>
<td><strong>ART vs observation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Todorokia 2000</td>
<td>Japan</td>
<td>1976–1999</td>
<td>pCCA</td>
<td>Na</td>
<td>ART</td>
<td>EBRT</td>
<td>17</td>
<td>N0:11; N1:2; N2:4</td>
<td>7 (41)</td>
<td>59.4 (34–76); 33.9; 32.0; 31.3;</td>
<td>–</td>
<td>10.0; 1.5</td>
<td></td>
</tr>
<tr>
<td>Jiang 2010</td>
<td>China</td>
<td>1998–2008</td>
<td>iCCA</td>
<td>–</td>
<td>ART</td>
<td>EBRT</td>
<td>24</td>
<td>N1: 8; Nc: 16</td>
<td>7 (29)</td>
<td>–</td>
<td>11.9; 19.1; 9.5; 1.5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Zheng 2018</td>
<td>China</td>
<td>2007–2016</td>
<td>iCCA</td>
<td>I–Na</td>
<td>ART</td>
<td>IMRT</td>
<td>26</td>
<td>N1: 23; Nc: 43</td>
<td>31 (47)</td>
<td>–</td>
<td>55.0; 19.1; 20.0; 10.0;</td>
<td>44.0; –</td>
<td>–</td>
</tr>
<tr>
<td><strong>ACRT vs observation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim 2011</td>
<td>Korea</td>
<td>2001–2009</td>
<td>eCCA</td>
<td>Iib, Iia, Iib</td>
<td>ACRT</td>
<td>5–FU/Leu/ Cap+ EBRT</td>
<td>75</td>
<td>N0: 39; N+: 36</td>
<td>17 (23)</td>
<td>60.0 (37–74); 41.1; 37.4;</td>
<td>–</td>
<td>34.8; 29.8; 21.6</td>
<td></td>
</tr>
<tr>
<td>Dover 2016</td>
<td>USA</td>
<td>2000–2013</td>
<td>CCA</td>
<td>I–IV</td>
<td>ACRT</td>
<td>GEM/5–FU+ EBRT</td>
<td>23</td>
<td>N0: 13; N+: 10</td>
<td>10 (43)</td>
<td>62.0 (38–80); 30.2; 26.3;</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Studies</td>
<td>Country 1</td>
<td>Study period</td>
<td>BTC type</td>
<td>Disease severity</td>
<td>Therapy</td>
<td>Regimen of CT/ type of RT</td>
<td>Sample size</td>
<td>Margin status (R0/R1)</td>
<td>Nodal status (negative/positive)</td>
<td>Female no (%)</td>
<td>Age, mean (SD/range), years</td>
<td>5-year OS (%); MOST (months)</td>
<td>5-year DFS (%); MDFST (months)</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>--------------</td>
<td>----------</td>
<td>------------------</td>
<td>---------</td>
<td>---------------------------</td>
<td>-------------</td>
<td>-----------------------</td>
<td>----------------------------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>ACRT vs ACT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gu 2017</td>
<td>China</td>
<td>2003–2013</td>
<td>GBC</td>
<td>II–Va</td>
<td>ACRT</td>
<td>Cap/S-1/Ox +5-FU/Cap/GEM</td>
<td>39</td>
<td>All R0</td>
<td>–</td>
<td>29 (74)</td>
<td>61 (35–77)</td>
<td>42.4; 27</td>
<td>48.7; 23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+ EBRT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>39</td>
<td></td>
<td>26 (67)</td>
<td>63 (45–85)</td>
<td>17.9; 13</td>
<td>13.5; 7</td>
<td></td>
</tr>
<tr>
<td>Hester 2018</td>
<td>USA</td>
<td>2004–2013</td>
<td>dCCA</td>
<td>I–IV</td>
<td>ACRT</td>
<td>–</td>
<td>348</td>
<td>258/76</td>
<td>N0:126; N+:211</td>
<td>135 (39)</td>
<td>64.3 (±10.0)</td>
<td>31.3 §; 32.1</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACT</td>
<td>348</td>
<td>256/78</td>
<td>N0:118; N+:206</td>
<td>136 (39)</td>
<td>64.4 (±10.8)</td>
<td>29.4 §; 29.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACT</td>
<td>20</td>
<td>16/2 (Rx:2)</td>
<td>17 (85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim 2016</td>
<td>Korea</td>
<td>2000–2015</td>
<td>GBC</td>
<td>I–IV</td>
<td>ACT GEM/Cis/OTHs</td>
<td>61</td>
<td>54/7</td>
<td>N0:21; N1/ Nx:36/4</td>
<td>42 (69)</td>
<td>64.5 (56–73)</td>
<td>Total: 33.0 §; 19.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACRT</td>
<td>–</td>
<td>44</td>
<td>36/8</td>
<td>28 (64)</td>
<td>63.9 (55–70)</td>
<td>Total: 74.6 §; 28.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>–</td>
<td>186</td>
<td>159/25</td>
<td>122 (66)</td>
<td>67.7 (56–74)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Im 2021</td>
<td>Korea</td>
<td>2001–2017</td>
<td>pCCA</td>
<td>II–Va</td>
<td>ACT 5–FU/GEM based</td>
<td>45</td>
<td>31/14</td>
<td>–</td>
<td>22 (33)</td>
<td>65 (32–82)†</td>
<td>38.9; –</td>
<td>24.6; –</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ART EBRT/IMRT</td>
<td>16</td>
<td>2/14</td>
<td>10 (63)</td>
<td>49.4; –</td>
<td>40.4; –</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACRT 5–FU/GEM based + EBRT/IMRT</td>
<td>16</td>
<td>5/11</td>
<td>4 (19)</td>
<td>49.4; –</td>
<td>40.4; –</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>–</td>
<td>37</td>
<td>25/12</td>
<td>35 (38.9)</td>
<td>27.0; –</td>
<td>24.5; –</td>
<td></td>
</tr>
<tr>
<td>Wan 2021</td>
<td>USA (SEER database)</td>
<td>1973–2015</td>
<td>GBC</td>
<td>II–IV</td>
<td>ACT</td>
<td>–</td>
<td>444</td>
<td>–</td>
<td>N0: 162; NxC: 282</td>
<td>311 (70)</td>
<td>–*</td>
<td>38.9/15.0; –</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACRT</td>
<td>542</td>
<td></td>
<td>N0: 156; NxC: 386</td>
<td>374 (69)</td>
<td>41.5/21.1; –</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>1703</td>
<td></td>
<td>N0: 961; NxC: 742</td>
<td>1192 (70)</td>
<td>32.8/11.3; –</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Characteristics of included RCT studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebata 2018</td>
<td>Japan</td>
<td>2007–2011</td>
<td>pCCA/ dCCA</td>
<td>I–III (T1–T4) ACT</td>
<td>GEM</td>
<td>–</td>
<td>117</td>
<td>106/11</td>
<td>N0: 75; N1: 42</td>
<td>40 (34)</td>
<td>–*</td>
<td>51.7; 62.3</td>
<td>45.0; 36.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>108</td>
<td>94/14</td>
<td>N0: 72; N1: 36</td>
<td>26 (24)</td>
<td>51.6; 63.8</td>
<td>44.0; 39.9</td>
<td></td>
</tr>
<tr>
<td>Edeline 2019</td>
<td>France</td>
<td>2009–2014</td>
<td>CCA/GBC</td>
<td>I–IV (T1–T4) ACT</td>
<td>GEMOX</td>
<td>–</td>
<td>95</td>
<td>82/13</td>
<td>N0:49; N+:35; Nx:11</td>
<td>38 (40)</td>
<td>63.0 (33–83)†</td>
<td>60.0 ± 75.8</td>
<td>–; 30.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Observation</td>
<td>99</td>
<td>87/12</td>
<td>N0:48; N:+36; N:15</td>
<td>49 (50)</td>
<td>63.0 (40–80)†</td>
<td>65.0 ± 50.8</td>
<td>–; 18.5</td>
</tr>
</tbody>
</table>
Table 1  Continued

<table>
<thead>
<tr>
<th>Studies</th>
<th>Country</th>
<th>Study period</th>
<th>BTC type</th>
<th>Disease severity</th>
<th>Therapy</th>
<th>Regimen of CT/ type of RT</th>
<th>Sample size</th>
<th>Margin status (R0/R1)</th>
<th>Nodal status (negative/positive)</th>
<th>Female no (%)</th>
<th>Age, mean (SD/range), years</th>
<th>5-year OS (%); MOST (months)</th>
<th>5-year DFS (%); MDFST (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primrose 2019†</td>
<td>UK</td>
<td>2006–2014</td>
<td>BTC</td>
<td>I–IV</td>
<td>ACT</td>
<td>Cap</td>
<td>223</td>
<td>N0: 121; N1: 102</td>
<td>112 (50)</td>
<td>62.0 (55–68)</td>
<td>25.1; 51.1</td>
<td>24.4</td>
<td></td>
</tr>
<tr>
<td>Observation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>224</td>
<td>N0: 115; N1:108</td>
<td>112 (50)</td>
<td>64.0 (55–69)</td>
<td>20.5; 36.4</td>
<td>17.5</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates that the data was not given originally. †Indicates that studies were presented as younger or older than a specific age. ‡Indicates that nodal status was given instead of mean age. §Indicates that margin status was given instead of mean age. ¶Indicates that the data was not given originally. **Indicates that studies were presented as younger or older than a specific age. ††Indicates that median age was given instead of mean age. †‡Indicates that survival rates were given instead of survival probability.

ACRT, adjuvant chemoradiation therapy; ACT, adjuvant chemotherapy; ART, adjuvant radiotherapy; BTC, biliary tract cancer; Cap, capecitabine; CCA, cholangiocarcinoma; Cis, cisplatin; CT, chemotherapy; dCCA, distal cholangiocarcinoma; 3DCRT, three-dimensional conformal radiotherapy; EBRRT, external beam radiotherapy; IMRT, intensity-modulated radiotherapy; Leu, Leucovorin; MDFST, median DFS time; MOST, median OS time; N+, positive lymph node; N+, unclear; OTHs, others; Ox, oxaliplatin; pCCA, perihilar cholangiocarcinoma (or hepatic duct carcinoma=HDC); R0, negative resection margin; R1, microscopic positive resection margin; RCT, randomised clinical trial; RT, radiotherapy; Rx, unknown; S-1, tegafur gimeracil oteracil potassium; Total, means the value of all patients in the study; 5-year DFS (%), 5-year disease-free survival probability; 5-year OS (%), 5-year overall survival probability.
were 46.17 for OS and 29.93 for DFS, which were similar to the DICs of the inconsistency model (46.05 and 31.15, respectively), suggesting no evidence of inconsistency in the network. The more parsimonious model, the consistency model, was used for our analyses (online supplemental table 8).

Network meta-analysis of treatments in BTC

A network meta-analysis was conducted to assess the efficacy (OS and DFS) of the following treatments in a Bayesian framework: observation alone after surgery, surgery with ACT, surgery with ART, and surgery with ACRT.

A network plot is shown in figure 2. OS data were available from 18 studies that included 5497 patients, of whom 2785 (50.7%) were in the observation group, 1519 (27.6%) received ACT, 83 (1.5%) received ART and 1110 (20.2%) received ACRT. The pooled OS data indicated that ART (HR 0.62; 95% CI 0.42 to 0.93), ACRT (HR 0.71; 95% CI 0.54 to 0.93) and ACT (HR 0.84; 95% CI 0.68 to 0.98) were more beneficial in patients with BTC compared with that of observation (figure 3A–B). No significant benefits in OS were observed in the comparisons between the different adjuvant therapies (figure 3A).

In terms of DFS, 14 studies with 1979 patients were included in the network meta-analysis (figure 2). The pooled DFS data demonstrated a significant improvement for ACRT and ACT compared with that of observation (HR 0.60; 95% CI 0.45 to 0.75, and HR 0.82; 95% CI 0.68 to 0.97, respectively). Furthermore, a slightly better efficacy for ACRT was obtained compared with that of ACT (HR 0.73; 95% CI 0.53 to 0.95). Significant differences in DFS were not observed between the other pairwise comparisons (figure 3A,C).

The ranking analysis was performed using SUCRA. Based on the pooled OS and DFS data, the ranking order of OS and DFS was inconsistent. The best therapy for OS was ranked as follows: ART, ACRT, ACT and observation (figure 3D). The best SUCRA value of ART was 87.0%, which was close to that of ACRT with a SUCRA value of 75.3% (online supplemental table 9). As for DFS, the best therapy was ranked as ACRT, ACT, ART, and observation (figure 3D). The SUCRA value of ACRT was approximately 97.1%, which was far higher than that of the others (online supplemental table 9).

Subgroup analyses for OS in patients with different residual tumour status, tumour sites, regions and absence of distant metastasis were performed. Seven studies...
reported the outcomes of patients after R0 resection, and seven studies reported the outcomes of patients after R1 resection (online supplemental figure 3A). In the R0 group, only ACRT had a survival advantage compared with that of surgery alone. In the R1 group, no survival advantage was observed in patients who underwent adjuvant therapies than in those with surgery alone (figure 4A, online supplemental figure 4A). In the subgroup analysis by primary tumour sites, 6 studies enrolled patients with GBC, and 11 studies recruited patients with CCA (online supplemental figure 3B). The OS benefit of ACRT was clear compared with that of observation in the CCA group (HR 0.62; 95% CI 0.36 to 0.85). The benefit of ART is unclear due to the lack of eligible ART studies. Among the comparable treatments in the GBC group, no significant differences in OS were found between ACT, ACRT and observation (figure 4B, online supplemental figure 4B). When studies were grouped according to region, 12 studies included patients in Asian countries and 7 from Western countries (online supplemental figure 3C). In both the Asian and Western groups, the pooled OS results favoured ACRT (HR 0.42; 95% CI 0.29 to 0.67 in Asia; HR 0.76; 95% CI 0.63 to 0.91 in Western countries) (figure 4C, online supplemental figure 4C). Subgroup analysis was also conducted for 15 studies investigating OS in patients with non-distant metastasis (online supplemental figure 3D). The pooled results showed superior efficacy of ACRT compared with that of observation (HR 0.74; 95% CI 0.60 to 0.85). Moreover, ACRT tended to be more effective compared with that of ACT (HR 0.82; 95% CI 0.67 to 0.96) (figure 4D, online supplemental figure 4D).

Subgroup analyses were also performed for DFS in patients with different tumour sites and regions and the absence of distant metastasis (online supplemental figure 5). Due to the small number of studies, subgroup analyses for DFS by different residual tumour statuses cannot be conducted. In patients without distant metastasis, the results from this subgroup were similar to that of the primary analysis; ACRT demonstrated statistically significant improvement with an HR of 0.61 (figure 4E, online supplemental figure 4D). When studies were split by primary tumour sites, only ACRT showed an apparent advantage in patients with CCA (figure 4F, online supplemental figure 4E). Furthermore, the stratified meta-analysis by region indicated that the favourable treatment was ACRT in Asia (HR 0.47; 95% CI 0.28 to 0.68) (figure 4G, online supplemental figure 4F).

To assess the robustness of the primary results, sensitivity analyses were performed for OS and DFS by excluding RCT studies, removing moderate-quality observational studies, and only including the studies for which HRs were reported in the original articles. The first sensitivity analysis included 15 retrospective studies of OS and 11 retrospective studies of DFS. The pooled results (online supplemental figure 6) and the ranking profiles of comparable treatments from retrospective studies confirmed the reliability of the primary OS results (figure 5A, online supplemental table 10). In terms of DFS, ACRT was still ranked as the best treatment option (figure 5B). Furthermore, the second and third sensitivity analyses were performed by removing moderate-quality observational studies (14 studies for OS and 11 retrospective studies of DFS). The results did not suggest a material change in the efficacy estimation for ACRT, but ART no longer showed advantages over observation in both OS and DFS; the improvement...
in OS and DFS between ACT and observation was insignificant except for the pooled DFS in high-quality studies (online supplemental figure 6).

The visual examination of the comparison-adjusted funnel plots did not suggest a publication bias for OS presented in our network meta-analysis (online supplemental figure 7A). The result of Egger’s test (p=0.251) also rejected the presence of small-study effects. However, both the visual examination of the comparison-adjusted funnel plots and the results of Egger’s test (p=0.018) suggested a publication bias for DFS in our network meta-analysis (online supplemental figure 7B).

**DISCUSSION**

BTCs are an uncommon and heterogeneous type of cancer with a higher prevalence in Asian countries. In general, BTCs include cancers raised from the intrahepatic, hilar, and distal bile ducts, as well as the gallbladder. Surgical resection provides the only chance for cure in patients with BTC at an early stage, but the survival outcomes are poor. The 5-year survival rate was as low as 10%. The most recent NCCN and ASCO guidelines recommend the use of adjuvant therapy for BTC patients after resection. However, experts mentioned that the use of adjuvant therapy is based on a very limited number of studies, and the benefit of adjuvant therapy remains unclear in many pivotal BTC trials. Therefore, it is necessary to compile up-to-date studies to validate the efficacy of adjuvant therapy in BTC patients.

In this network meta-analysis, 19 studies were included to evaluate the comparative benefits (involving 5497 patients for OS and 1979 patients for DFS) between adjuvant therapy (ACT, ART and ACRT) following surgical resection and curative-intent resection (observation group). Our primary results demonstrated that adjuvant therapy was more effective than that of observation in OS. However, no statistically significant difference was detected between ACT, ACRT and ART. Moreover, the pooled DFS results suggested a statistically significant benefit for ACRT and ACT over observation. Although ART was ranked first in OS, its DFS result was inconsistent with OS, and caution should be exercised regarding the bias that arises from the small sample size effect of ART evidence (only 83 patients). In addition, a previously published meta-analysis of adjuvant therapy in the treatment of BTC suggested that ART did not provide a significant advantage over observation. We hold reservations on the conclusion of the efficacy of ART, and further evidence is needed to elucidate this matter.

Combining the results of OS with DFS, ACRT, and ACT after radical resection could provide a survival benefit in patients with BTC. This was in line with another meta-analysis that showed that both ACT (HR 0.61; 95% CI 0.47 to 0.79) and ACRT (HR 0.35; 95% CI 0.14 to 0.83) could significantly improve the clinical survival of patients with resected BTC compared with that of surgery alone. As known, the BILCAP study is a unique positive randomised trial of ACT in patients with BTC. Despite the concerns regarding the findings of the BILCAP trial, this study established adjuvant capecitabine as the new standard of care for resected BTC. However, no randomised trials of ACRT vs ACT are available. Although in our study, the ACRT ranked high in terms of OS and DFS (second place for OS, first place for DFS), whether it is genuinely superior to ACT needs further confirmation in prospective RCTs.

The types of patients who are more likely to benefit from adjuvant therapy are also the focus. A previous phase II study indicated that the risk factors were GBC, eCCA, pathological stage T2–4, positive LN or positive RM. In our study, we were only able to conduct subgroup analyses on the effects of the RM, primary tumour site, regions and absence of distant metastasis, but not LN and tumour size, on adjuvant therapy due to the availability of data. We observed that ACRT showed a modest improvement in DFS than that of ACT in CCA and Asian patients and ranked first in each subgroup (online supplemental table 11). We noticed that approximately 95% (for OS analysis) and 80% (for DFS analysis) of CCA patients treated with ACRT had eCCA. The benefit of ACRT observed in CCA patients may be mainly derived from eCCA patients. Inconsistent with most studies, ACRT has an OS benefit in the R0 group, but not in the R1 group. This was possibly due to the more stringent selection of studies and the effects of small sample sizes. Subgroup analyses for DFS by R0 or R1 were not conducted because of limited data. Therefore, the subgroups of patients who could benefit more from postoperative adjuvant therapy need to be further elucidated. We look forward to seeing more randomised controlled studies in this field, especially head-to-head trials and trials designed for more specific subgroups. Using the primary tumour site, tumour size, disease severity, LN metastases and RM as stratification factors in future studies would reduce selection bias due to the heterogeneity of the population.


---

**Figure 5** Bayesian ranking plots of comparable treatments on efficacy in patients with biliary tract cancer in the sensitivity analyses. Ranking curves indicate the probability of each comparable treatment being ranked from first to last on overall survival (OS) (solid lines) and disease-free survival (DFS) (dotted lines). Data of the curves are presented in online supplemental table 10. ACT, adjuvant chemotherapy; ACRT, adjuvant chemoradiation therapy; ART, adjuvant radiotherapy.
This study had several limitations. First, RCT data were only available to compare ACT and observations. Unfortunately, there are neither RCTs for ART and ACRT nor head-to-head RCTs between different adjuvant treatments. The notable differences in the study design level may introduce confounding factors in our data analysis, although data transitivity and consistency could be assumed statistically. Second, the included studies spanned over a 45-year period during which operative techniques and methods have changed and improved over time. These changes in treatment methods could potentially bias our results, but the impact on the outcome was unclear and difficult to interpret. Third, most comparisons were indirect, and direct evidence was obtained from two or three studies. The comparisons between the different treatment modalities and treatment regimens may substantially contribute to heterogeneity among included studies. Patients in ACT, ART or ACRT groups were treated with different modalities. Even within the same treatment modality, different regimens offered various efficacy results. Such as, the adjuvant capcitabine monotherapy (BILCAP study) appeared more effective compared with observation, whereas adjuvant gemcitabine (BCAT study) or gemcitabine plus oxaliplatin (PRODIGE 12 study) did not.98 Furthermore, we noticed that the number of patients treated with ART was small. These may result in a considerable risk of bias. Fourth, the definitions of OS and DFS in RCTs and retrospective studies were calculated differently. In RCTs, it started from the date of randomisation to the date of surgery in retrospective studies. The span was relatively short compared with the expected survival duration, but the slight variation in data collection should still be considered. Finally, we failed to compare the safety outcomes due to a lack of sufficient data on adverse events. Taking all these into account, our estimates should be interpreted with caution.

CONCLUSIONS

Our primary results demonstrated that, compared with that of observation, ACRT and ACT after radical resection could provide better OS and DFS benefits in patients withBTC. However, ART only showed improvement in OS. ACRT had a modest DFS advantage compared with that of ACT. Due to the absence of direct evidence from head-to-head prospective studies, thorough and high-quality RCTs are warranted to consolidate our results further. Optimal regimens and dosing schedules still need to be explored.

Acknowledgements We thank the study authors Zhaochong Zeng (Jiang, 2010), Tae Hyun Kim (Kim, 2013), Wei-Hu Wang (Zheng, 2018) and Xiaojian. Ni (Wan, 2021) responded to our request for additional data and questions. We thank Lei Wang for the help in the manuscript proofreading.

Contributors As the guarantor, MO designed, coordinated and supervised this study. YCh, BZ and CLi screened the articles, YCh and BZ extracted the data and assessed the quality. YCa and CLi contacted study authors for additional information. YCh, BZ and YCa analysed and prepared the material. YCh, BZ, CLi and CLu interpreted the data and wrote the draft of the report. All authors had reviewed and approved the final submitted version of the report. MO was responsible for the integrity and accuracy of the data.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information. All data relevant to the study are included in the article or uploaded as supplemental information. All the data extracted, and analysed, are reported in this paper were from published literature.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Baoxia Zhang http://orcid.org/0000-0003-2445-3031

REFERENCES

Begg CB, Mazumdar M. Operating characteristics of a RANK detected by a simple, graphical test. *Int J Epidemiol* 2010;29:932–44.


