A cumulative meta-analysis of endoscopic papillary balloon dilation versus endoscopic sphincterotomy for removal of common bile duct stones

Authors
Alberto Tringali1, Matteo Rota2,3, Marta Rossi2, Cesare Hassan4, Douglas G. Adler5, Massimiliano Mutignani1

Institutions
1 Endoscopy Unit, Department of Surgery, Ospedale Niguarda Ca Granda, Milan, Italy
2 Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy
3 Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy
4 Endoscopy Unit, Nuovo Regina Margherita Hospital, Rome, Italy
5 Division of Gastroenterology and Hepatology, University of Utah Hospital, Salt Lake City, Utah, United States

submitted 7.4.2018
accepted after revision 25.10.2018

Bibliography
DOI https://doi.org/10.1055/a-0818-3638
Published online: 2019 | Endoscopy
© Georg Thieme Verlag KG Stuttgart · New York
ISSN 0013-726X

Corresponding author
Alberto Tringali, MD. Endoscopy Unit, Department of Surgery, Niguarda Hospital, Piazza Ospedale Maggiore 3, 20162 Milan, Italy
Fax: +39-02-64442911
albri10@gmail.com
alberto.tringali@ospedaleniguarda.it

Fig. e1, e2, e4, e6, e8, e10, Table e3 – e7, Supplementary material
Online content viewable at: https://doi.org/10.1055/a-0818-3638

ABSTRACT

Background Endoscopic papillary balloon dilation (EPBD) was introduced to overcome the risk of adverse events associated with endoscopic sphincterotomy in the removal of common bile duct (CBD) stones. We performed a meta-analysis of randomized controlled trials (RCTs) comparing efficacy and safety of EPBD vs. endoscopic sphincterotomy, focusing on stone size, balloon diameter, and balloon dilation time.

Methods A multiple database search was performed, including MEDLINE, EMBASE and Cochrane Library, from their inception date until October 2017. RCTs comparing the efficacy and safety of EPBD vs. endoscopic sphincterotomy in the removal of CBD stones were included. Cumulative meta-analyses over time, and subgroup analyses according to stone size, and balloon diameter and dilation time were carried out.

Results 25 RCTs met the inclusion criteria. Despite the cumulative meta-analysis showing a trend over time in favor of endoscopic sphincterotomy in studies published up to 2004, the conventional meta-analysis revealed that EPBD was equally efficacious compared with endoscopic sphincterotomy in stone removal at first attempt (odds ratio [OR] 0.95, 95% confidence interval [CI] 0.65 – 1.38). Endoscopic sphincterotomy was superior to EPBD in terms of overall stone clearance (OR 0.65, 95% CI 0.43 – 0.99) in studies published since 2002, but no differences emerged in studies using large (≥10 mm) balloons (OR 1.37, 95% CI 0.72 – 2.62). No statistically significant difference in pancreatitis occurrence emerged between EPBD and endoscopic sphincterotomy (OR 1.35, 95% CI 0.90 – 2.03). Pancreatitis was more common with EPBD than with endoscopic sphincterotomy in studies using balloons <10 mm (OR 1.78, 95% CI 1.07 – 2.97), whereas no difference emerged in studies using large balloons (OR 0.84, 95% CI 0.46 – 1.53). EPBD had lower rates of bleeding and cholecystitis.

Conclusions Our latest data confirm that EPBD is currently inferior to endoscopic sphincterotomy in terms of overall stone clearance. However, EPBD using large balloons (≥10 mm) was as effective as endoscopic sphincterotomy, both in stone clearance and the need for endoscopic mechanical lithotripsy, without carrying an increased risk of pancreatitis.
Introduction

Endoscopic papillary balloon dilation (EPBD) is an alternative to endoscopic sphincterotomy for the removal of bile duct stones in the context of endoscopic retrograde cholangiopancreatography (ERCP). EPBD, which was first proposed in 1983 by Staritz et al. [1], has several potential advantages over endoscopic sphincterotomy. Primarily, EPBD avoids cutting the biliary sphincter, thus reducing early adverse events such as bleeding and perforation. Furthermore, it may be a reasonable option in patients with coagulopathy or surgically altered anatomy who cannot undergo sphincterotomy [2,3]. Despite the potential advantages, the significantly higher reported risk of pancreatitis in the literature when this maneuver is performed [4–7] has led many to prefer endoscopic sphincterotomy over EPBD [8–9].

Several systematic reviews and meta-analyses of randomized controlled trials (RCTs) have compared the safety and efficacy of EPBD and endoscopic sphincterotomy [4,7,10–13]. These reviews generally found that EPBD was inferior to endoscopic sphincterotomy in terms of overall stone removal; however, it was associated with a lower risk of bleeding when compared with endoscopic sphincterotomy. The latter meta-analysis by Park et al. [13] also showed that there was no significant increased risk of post-ERCP pancreatitis (PEP) with EPBD compared with endoscopic sphincterotomy. This supports the hypothesis of some earlier papers suggesting that the risk of EPBD-associated PEP may be lower than originally thought [14–15].

However, these meta-analyses did not evaluate whether the safety and efficacy of EPBD vs. endoscopic sphincterotomy differed according to stone size, or according to the size of the balloon or its dilation time. A consensus guideline [16] for endoscopic papillary large (≥12 mm) balloon dilation (EPLBD) showed that EPLBD can be used as the initial method when a large bile duct stone (>15 mm) has been identified on ERCP or cross-sectional imaging [16].

The aim of this study, therefore, was to perform a cumulative meta-analysis of the results of RCTs comparing the efficacy and safety of EPBD vs. endoscopic sphincterotomy, with special focus on stone size, balloon diameter, and balloon dilation time.

Methods

This systematic review and meta-analysis was carried out according to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [17] (see Supplementary material 1).

Identification and selection of studies

Inclusion criteria were RCTs that compared the efficacy and safety of EPBD vs. endoscopic sphincterotomy for removal of common bile duct (CBD) stones found on ERCP in patients aged 18 years old or older.

Studies were excluded if they were nonrandomized studies, studies analyzing post-sphincterotomy balloon dilation, and studies considering patients who had undergone previous treatment for CBD stones.

Search strategy

A literature search was performed in EMBASE, MEDLINE, the Cochrane library, and the Cochrane Central Register of Controlled Trials (CENTRAL) from their inception date to October 2017 (see Supplementary material 2 for search strategy). For the sake of completeness, we searched references of relevant literature. After removal of duplicates, titles and abstracts were screened by two authors (A. Tringali and M. Rota) to identify potentially eligible studies. Differences were resolved by discussion.

Data collection

Data were extracted using a standardized form. The following data were extracted:

- Descriptive data – first author last name, year of publication, country of origin (grouped into Western and Asian), study setting, number of patients, age and sex distribution of patients, previous ERCP, altered anatomy (e.g. Billroth II), presence of perihepatic diverticulum, stone size, stone number, CBD size, mean follow-up time, balloon diameter, duration of balloon dilatation, pressure of balloon dilatation.

Outcome data – primary outcome measures were stone clearance at the first session and overall stone clearance considering subsequent stone extraction attempts. Secondary outcome measures included the need for endoscopic mechanical lithotripsy (EML), pancreatitis, bleeding, perforation, cholangitis, cholecystitis, recurrence of CBD stones, 30-day procedure mortality, overall short-term procedure-related complication rate (defined as composite outcomes of all complications), and procedure time (defined as elapsed time from ERCP to the end of the procedure).

- Qualitative data – random sequence generation, allocation concealment, blinding of participants and personnel, blinding outcome assessment, incomplete outcome data, selective reporting, and loss to follow-up.

Quality appraisal

Each included study was appraised for quality by two independent evaluators (A. Tringali and M. Rota). Quality appraisal was performed using the risk of bias tool as recommended by the Cochrane collaboration [18].

Statistical analysis

Dichotomous outcomes were evaluated in terms of crude odds ratio (ORs) with their 95% confidence intervals (CIs), and summarized across studies through a random-effects model. Briefly, each study-specific log (OR) was weighted by the inverse of its variance plus an estimate of the between-study variance component τ² computed through the DerSimonian and Laird moment estimator [19]. If no between-study heterogeneity was evident – resulting in a null τ² estimate – the pooled estimator from the random-effects model would be equal to the one derived from a fixed-effects model. The procedure time
was analyzed using the Hedges’ standardized mean difference estimator.

The 95% prediction intervals, which estimate the expected effect in future studies, were also computed.

Cumulative meta-analyses were performed to investigate the effect of publication year on the pooled estimates. Briefly, in a cumulative meta-analysis, studies were added one at a time according to publication year (earliest to the most recent), and results were displayed in a forest plot showing the pooled estimate each time the results of a new study were added over time.

Between-studies heterogeneity was assessed through the Q test based on the chi-squared statistics, and inconsistency was quantified through the $I^2$ statistic [20]. In order to assess potential sources of heterogeneity, we first performed a sensitivity analysis by removing each study in turn (leave-one-out method) to evaluate its influence on the final pooled estimate.

Stratified analyses were carried out according to the geographic area of the study (Asian vs. Western), balloon diameter (<10 vs. ≥10 mm), and dilation time (≤1 vs. >1 minute), and according to the largest CBD stone size (<10 vs. ≥10 mm) when reported in the original publications.

Publication bias was assessed by visual inspection of funnel plots for asymmetry and through the Egger’s test for asymmetry [21].

Data were synthesized using the “metafor” library of R version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Study selection

A total of 1267 unique studies were identified through the systematic review of the literature. Following screening of abstracts and titles, we identified 59 potentially eligible studies for which full-text reading was required. Finally, 25 articles [2, 5, 22–44] were included (Fig.e1, available online).

Characteristics of the included studies

The main characteristics of the 25 included studies [2, 5, 22–44] are reported in Table 1. The studies were published between 1995 and 2017, and included a total of 3360 patients, 1665 of whom were randomized to EPBD and 1695 to endoscopic sphincterotomy. The majority of the studies (n=18) were performed in Asian countries [22, 25–27, 29–31, 33–43], and the remaining ones [2, 5, 23–24, 28, 32, 44] were conducted in Western countries (the Netherlands, USA, Germany, UK, and Egypt). Seven studies only used balloons that were ≥10 mm in diameter [32, 38–40, 42–44], whereas the majority of the remaining studies used a balloon with a diameter of 8 mm [2, 5, 22–25, 27–31, 33–37, 41]. The balloon dilation lasted for more than 1 minute in half of the studies [22, 25, 27–30, 33–36, 39–41].

The risk of bias evaluation according to the Cochrane risk of bias tool did not show significant bias (Fig.e2, available online).

Primary outcomes

A total of 18 studies, including 1262 EPBD and 1284 endoscopic sphincterotomy patients, reported the success rate of stone clearance at the first session (Table 2). There was no statistically significant difference between EPBD and endoscopic sphincterotomy (OR 0.95, 95%CI 0.65–1.38), with no single study influencing the pooled estimate. However, the cumulative meta-analysis showed a statistically significant change over time ($P<0.01$) in the success rate of stone clearance at the first session (Fig.3a). In detail, the OR for studies published up to 2004 (n=11) was 0.63 (95%CI 0.45–0.87), showing that endoscopic sphincterotomy was superior to EPBD at the first session. Conversely, the OR for studies published after 2004 (n=7) was 1.97 (95%CI 1.31–2.97), showing a two-fold increased success rate at the first session for patients undergoing EPBD compared with those undergoing endoscopic sphincterotomy.

In terms of overall stone clearance (Table 2), endoscopic sphincterotomy was superior to EPBD (OR 0.65, 95%CI 0.43–0.99). The cumulative meta-analysis showed that endoscopic sphincterotomy performed better than EPBD starting from 2002, although the change in overall stone clearance over time was not significant ($P=0.13$) between the two endoscopic techniques (Fig.3b).

Secondary outcomes

The results of secondary outcomes are provided in Table 2 (see Table e3, available online, for study-specific data). There were 23 studies, involving 1507 EPBD and 1531 endoscopic sphincterotomy patients, reporting the need for EML for CBD stone removal. The pooled OR was 1.24 (95%CI 1.09–1.40), showing that the need for EML was slightly more common in the EPBD group compared with the endoscopic sphincterotomy group (Fig.e4, available online). The leave-one-out sensitivity analysis showed that the pooled estimate was statistically significant after the exclusion of the study by Kogure et al. [43]. The cumulative meta-analysis (Fig.5) showed that the increased need for EML for EPBD vs. endoscopic sphincterotomy significantly reduced over time ($P<0.01$). In fact, the need for EML in EPBD progressively decreased vs. endoscopic sphincterotomy, losing its statistical significance beyond 2015.

A total of 22 studies, including a total of 1420 EPBD and 1448 endoscopic sphincterotomy patients, reported PEP data. There was a slightly, but not statistically significant, increased odds of PEP in EPBD vs. endoscopic sphincterotomy patients (OR 1.35, 95%CI 0.90–2.03; Fig.e6, available online). The leave-one-out sensitivity analysis showed that the odds of PEP was statistically significantly higher (OR 1.52, 95%CI 1.05–2.18) after the exclusion of the Fu et al. study [39]. The cumulative meta-analysis (Fig.7a) did not show a statistically significant change over time ($P=0.23$), although the odds of PEP was lower for studies published before 2003 (OR 1.17, 95%CI 0.65–2.10) compared with those published after (OR 1.49, 95%CI 0.81–2.75).
### Table 1 Main characteristics of the studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study and country</th>
<th>Patient characteristics</th>
<th>Balloon characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients, n</td>
<td>Age, years</td>
</tr>
<tr>
<td>Minami et al., 1995 [22] Japan (Asian)</td>
<td>EPBD 20</td>
<td>64 ± 11.2</td>
</tr>
<tr>
<td>Bergman et al., 1997 [23] The Netherlands (Western)</td>
<td>EPBD 101</td>
<td>72 (27 – 98)</td>
</tr>
<tr>
<td>Chen et al., 1998 [24] USA (Western)</td>
<td>EPBD 47</td>
<td>52 ± 20</td>
</tr>
<tr>
<td>Cho et al., 1998 [25] Korea (Asian)</td>
<td>EPBD 42</td>
<td>58.9</td>
</tr>
<tr>
<td>Iwata et al., 1998 [26] Japan (Asian)</td>
<td>EPBD 37</td>
<td>N/A</td>
</tr>
<tr>
<td>Ochi et al., 1999 [27] Japan (Asian)</td>
<td>EPBD 55</td>
<td>62.6 ± 15.9</td>
</tr>
<tr>
<td>Arnold et al., 2001 [28] Germany (Western)</td>
<td>EPBD 30</td>
<td>54.2 ± 18.5</td>
</tr>
<tr>
<td>Bergman et al., 2001 [2] The Netherlands (Western)</td>
<td>EPBD 16</td>
<td>73 (43 – 84)</td>
</tr>
<tr>
<td>Yasuda et al., 2001 [29] Japan (Asian)</td>
<td>EPBD 35</td>
<td>69.5 (42 – 86)</td>
</tr>
<tr>
<td>Natsui et al., 2002 [30] Japan (Asian)</td>
<td>EPBD 70</td>
<td>64.5 (23 – 87)</td>
</tr>
<tr>
<td>Fujita et al., 2003 [31] Japan (Asian)</td>
<td>EPBD 138</td>
<td>66.8 (26 – 93)</td>
</tr>
<tr>
<td>Vlavianos et al., 2003 [32] UK (Western)</td>
<td>EPBD 103</td>
<td>60.8 ± 20.5</td>
</tr>
<tr>
<td>Disario et al., 2004 [5] USA (Western)</td>
<td>EPBD 117</td>
<td>47 ± 19</td>
</tr>
<tr>
<td>Lin et al., 2004 [33] Taiwan (Asian)</td>
<td>EPBD 51</td>
<td>64 (28 – 90)</td>
</tr>
<tr>
<td>Takezawa et al., 2004 [34] Japan (Asian)</td>
<td>EPBD 46</td>
<td>70 (40 – 90)</td>
</tr>
</tbody>
</table>
The odds of bleeding were statistically significantly lower in EPBD vs. endoscopic sphincterotomy (OR 0.30, 95% CI 0.16–0.55). Although recurrence of CBD stones was less common after EPBD than after endoscopic sphincterotomy, there was no statistically significant difference (OR 0.77, 95% CI 0.52–1.13).

EPBD had a lower occurrence of short-term cholecystitis than endoscopic sphincterotomy (OR 0.35, 95% CI 0.13–0.93). A total of 23 studies, including a total of 1477 EPBD and 1505 endoscopic sphincterotomy patients, reported overall short-term complications. There was no statistically significant difference in the incidence of overall short-term complications (OR 0.94, 95% CI 0.64–1.36) in EPBD vs. endoscopic sphincterotomy (Fig. e8, available online). The cumulative meta-analysis (Fig. 7b) did not show a statistically significant difference (P=0.73) over the time span from the first considered study published in 1995 [22] to the latest one published in 2017 [44].

**Subgroup analyses**

The main results for subgroup analyses are reported in Fig. 9 (see also Table 4, Table 5, Table 6, Table 7, available online). In terms of overall stone clearance (Fig. 9a), endoscopic sphincterotomy was superior to EPBD in Asian (OR 0.45, 95% CI 0.24–0.84) but not in Western studies (OR 0.81, 95% CI 0.40–1.66), in studies using a balloon with a diameter <10 mm (OR 0.42, 95% CI 0.25–0.72), and in studies where the balloon was dilated for more than 1 minute (OR 0.34, 95% CI 0.18–0.67).
Table 2: Pooled meta-analytic odds ratios for the comparison of endoscopic papillary balloon dilation vs. endoscopic sphincterotomy for bile duct stone removal.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of studies</th>
<th>Events/Patients, n EPBD vs. EST</th>
<th>OR (95%CI) I², % Prediction interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stone clearance at the first session</td>
<td>18</td>
<td>1013/1262 vs. 1034/1284</td>
<td>0.95 (0.65–1.38) 63¹ 0.25–3.67</td>
</tr>
<tr>
<td>Overall stone clearance</td>
<td>20</td>
<td>1109/1185 vs. 1157/1202</td>
<td>0.65 (0.43–0.99)³ 3 0.37–1.14</td>
</tr>
<tr>
<td><strong>Secondary outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for EML</td>
<td>23</td>
<td>309/1507 vs. 271/1531</td>
<td>1.24 (0.92–1.67) 41¹ 0.47–3.25</td>
</tr>
<tr>
<td>PEP</td>
<td>22</td>
<td>105/1420 vs. 72/1448</td>
<td>1.35 (0.90–2.03) 24 0.46–3.93</td>
</tr>
<tr>
<td>Bleeding</td>
<td>19</td>
<td>8/1234 vs. 51/1264</td>
<td>0.30 (0.16–0.55)³ 0 0.15–0.57</td>
</tr>
<tr>
<td>Perforation</td>
<td>16</td>
<td>3/1113 vs. 5/1142</td>
<td>0.89 (0.37–2.13) 0 0.34–2.32</td>
</tr>
<tr>
<td>Short-term cholangitis</td>
<td>12</td>
<td>18/884 vs. 16/893</td>
<td>1.06 (0.51–2.18) 0 0.46–2.41</td>
</tr>
<tr>
<td>Long-term cholangitis</td>
<td>5</td>
<td>5/400 vs. 10/412</td>
<td>0.67 (0.22–2.06) 0 0.11–4.14</td>
</tr>
<tr>
<td>Short-term acute cholecystitis</td>
<td>6</td>
<td>4/460 vs. 17/474</td>
<td>0.35 (0.13–0.93)³ 0 0.09–1.39</td>
</tr>
<tr>
<td>Long-term acute cholecystitis</td>
<td>7</td>
<td>8/396 vs. 16/402</td>
<td>0.52 (0.22–1.26) 0 0.16–1.66</td>
</tr>
<tr>
<td>CBD stone recurrence</td>
<td>13</td>
<td>57/816 vs. 75/836</td>
<td>0.77 (0.52–1.13) 0 0.50–1.18</td>
</tr>
<tr>
<td>30-days procedure-related mort-</td>
<td>22</td>
<td>5/1441 vs. 2/1465</td>
<td>1.22 (0.55–2.68) 0 0.53–2.81</td>
</tr>
<tr>
<td>Jality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall short-term complications</td>
<td>23</td>
<td>165/1477 vs. 178/1505</td>
<td>0.94 (0.64–1.36) 52¹ 0.24–3.69</td>
</tr>
<tr>
<td>Procedure time</td>
<td>9</td>
<td>529 vs. 548</td>
<td>0.12 (0.07 to 0.20) 83¹ 0.67 to 0.92</td>
</tr>
</tbody>
</table>

CBD, common bile duct; CI, confidence interval; EML, endoscopic mechanical lithotripsy; EPBD, endoscopic papillary balloon dilation; EST, endoscopic sphincterotomy; OR, odds ratio; PEP, post-endoscopic retrograde cholangiopancreatography pancreatitis.

¹ Statistically significant at α=0.05.
² Standardized mean difference (95%CI)

There was no statistically significant difference in the need for EML (Fig. 9b) between EPBD and endoscopic sphincterotomy in both Asian (OR 1.26, 95%CI 0.91–1.76) and Western studies (OR 1.13, 95%CI 0.52–2.42). The need for EML was required less with endoscopic sphincterotomy than with EPBD in studies using a balloon with a diameter <10 mm (OR 1.72, 95% CI 1.33–2.23), whereas in studies using larger balloons, the need for EML was less frequent with EPBD than with endoscopic sphincterotomy (OR 0.62, 95%CI 0.42–0.91). Conversely, endoscopic sphincterotomy required less EML than EPBD in studies where the balloon was dilated for more than 1 minute (OR 1.66, 95%CI 1.21–2.27). Moreover, the need for EML was less frequent with endoscopic sphincterotomy than with EPBD (OR 2.66, 95%CI 1.31–5.40) for stones <10 mm, but no statistically significant differences emerged for stones ≥10 mm (OR 1.07, 95%CI 0.52–2.22).

A slightly higher occurrence of PEP (Fig. 9c) emerged for EPBD vs. endoscopic sphincterotomy in studies from Western countries (OR 2.33, 95%CI 1.00–5.42) but not in those from Asian countries (OR 1.12, 95%CI 0.72–1.76). The odds of PEP were statistically significantly higher for EPBD vs. endoscopic sphincterotomy in studies using a balloon with a diameter <10 mm (OR 1.78, 95%CI 1.07–2.97), whereas there was no statistically significant difference in PEP occurrence (OR 0.84, 95%CI 0.46–1.53) in studies using larger balloons. PEP was more common with EPBD (OR 2.29, 95%CI 0.93–5.62) than with endoscopic sphincterotomy in studies with a short (≤1 minute) balloon dilation, whereas no difference in PEP occurrence emerged in studies where the balloon was dilated for more than 1 minute (OR 1.12, 95%CI 0.68–1.85).

**Publication bias**

Visual inspection of funnel plots for overall stone clearance and overall short-term complications showed no evidence of asymmetry (Fig. e10, available online). As a confirmation, the Egger’s test for funnel plot asymmetry gave a P value of 0.49 and 0.61, respectively, showing no potential publication bias for the two considered outcomes.

**Discussion**

We carried out a cumulative meta-analysis to investigate how the efficacy and safety of EPBD vs. endoscopic sphincterotomy in CBD stone removal varied over time and to study, for the first time, the impact of balloon characteristics (e.g., balloon diameter and dilation time) and stone size on overall stone clearance,
Fig. 3 Forest plots showing the results of conventional and cumulative meta-analyses of endoscopic papillary balloon dilation (EPBD) vs. endoscopic sphincterotomy (EST) for the removal of common bile duct stones. a Stone clearance at the first session. b Overall stone clearance. In the conventional meta-analysis, the summary odds ratio (OR) is represented through a diamond and its tips represents the 95% confidence interval (CI), whereas the prediction interval is represented through a dashed thin line. In the cumulative meta-analysis, a summary OR was calculated each time the results of a new study were added. The rate of successful stone clearance at the first session was not significantly different between EPBD and endoscopic sphincterotomy starting from 2007, whereas endoscopic sphincterotomy was superior to EPBD in terms of overall stone clearance since 2002.
EML use, and PEP. With respect to the conventional approach, cumulative meta-analysis allowed us to study the differences between EPBD and endoscopic sphincterotomy over time. In particular, we were able to assess the impact of each new study on the pooled estimate, and thus to pinpoint the first time a difference in outcome becomes statistically significant.

In line with the latest published meta-analysis [13], our results, which importantly included additional RCTs [24–26, 40, 43–44] not previously considered, showed that there was an absence of evidence of a meaningful difference between EPBD and endoscopic sphincterotomy in terms of successful stone removal rates at the index procedure. Of note, the cumulative meta-analysis clearly evidenced a trend over time in favor of endoscopic sphincterotomy up to 2004, and thereafter no significant difference emerged between the two endoscopic techniques. As shown in studies published after 2004, it is possible that increasing experience with EPBD reduced the difference in stone removal rates at first attempt compared with endoscopic sphincterotomy. This finding could possibly be related to better selection of patients for endoscopic sphincterotomy and to a greater use of large-diameter (≥10 mm) balloons for EPBD after 2004, which showed an improved clearance at the index procedure compared with endoscopic sphincterotomy in previous studies [45–46], as in our investigation.

EPBD was inferior to endoscopic sphincterotomy in terms of overall stone clearance in studies published since 2002, and more frequently required EML probably because bile duct dilation may not be sufficient for CBD stone removal in all patients [13]. Notably, in studies using large (≥10 mm) balloons, EPBD appeared superior to endoscopic sphincterotomy in terms of overall stone clearance, requiring less EML use. Although only a few studies have reported data stratified by stone size, use of EML was more frequently required for removal of small (≤10 mm) stones in EPBD than with endoscopic sphincterotomy. This may be due to inadequate papillary dilation caused by the use of small balloons to potentially limit the risk of PEP.

Although PEP was more common with EPBD than with endoscopic sphincterotomy in studies published since 2003, our data confirmed an absence of evidence of a difference in PEP occurrence in the EPBD group compared with the endoscopic sphincterotomy group; this finding was also reported by Seo et al. [47], who suggested that papillary balloon dilation may not in itself be a cause of PEP. The development of PEP might be related to the trauma to the papilla following balloon dilation rather than to the balloon dilation itself [47].

Subgroup analyses revealed an increased, though not significant, PEP risk in studies where papillary dilation lasted for less than 1 minute. This in accordance with the results of an RCT.
<table>
<thead>
<tr>
<th>Study</th>
<th>EPBD</th>
<th>EST</th>
<th>OR [95 % -CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minami 1995</td>
<td>2</td>
<td>2</td>
<td>1.00 [0.13, 7.89]</td>
</tr>
<tr>
<td>+ Bergman 1997</td>
<td>9</td>
<td>121</td>
<td>1.00 [0.38, 2.61]</td>
</tr>
<tr>
<td>+ Chen 1998</td>
<td>12</td>
<td>168</td>
<td>1.18 [0.49, 2.87]</td>
</tr>
<tr>
<td>+ Iwata 1998</td>
<td>14</td>
<td>205</td>
<td>1.00 [0.45, 2.22]</td>
</tr>
<tr>
<td>+ Ochi 1999</td>
<td>14</td>
<td>256</td>
<td>0.91 [0.42, 1.95]</td>
</tr>
<tr>
<td>+ Arnold 2001</td>
<td>20</td>
<td>286</td>
<td>1.10 [0.56, 2.17]</td>
</tr>
<tr>
<td>+ Bergman 2001</td>
<td>21</td>
<td>302</td>
<td>1.15 [0.59, 2.38]</td>
</tr>
<tr>
<td>+ Yasuda 2001</td>
<td>23</td>
<td>337</td>
<td>1.14 [0.60, 2.14]</td>
</tr>
<tr>
<td>+ Natsui 2002</td>
<td>27</td>
<td>407</td>
<td>1.54 [0.91, 2.58]</td>
</tr>
<tr>
<td>+ Fujita 2003</td>
<td>42</td>
<td>545</td>
<td>1.64 [0.99, 2.72]</td>
</tr>
<tr>
<td>+ Vlavianos 2003</td>
<td>47</td>
<td>648</td>
<td>1.91 [1.05, 3.48]</td>
</tr>
<tr>
<td>+ DiSario 2004</td>
<td>65</td>
<td>765</td>
<td>1.88 [1.06, 3.34]</td>
</tr>
<tr>
<td>+ Lin 2004</td>
<td>65</td>
<td>816</td>
<td>1.96 [1.17, 3.26]</td>
</tr>
<tr>
<td>+ Tanaka 2004</td>
<td>68</td>
<td>832</td>
<td>1.84 [1.12, 3.02]</td>
</tr>
<tr>
<td>+ Watanabe 2007</td>
<td>77</td>
<td>922</td>
<td>1.54 [0.88, 2.68]</td>
</tr>
<tr>
<td>+ Oh 2012</td>
<td>79</td>
<td>962</td>
<td>1.51 [0.91, 2.50]</td>
</tr>
<tr>
<td>+ Fu 2013</td>
<td>82</td>
<td>1050</td>
<td>1.48 [0.92, 2.36]</td>
</tr>
<tr>
<td>+ Minakari 2013</td>
<td>91</td>
<td>1130</td>
<td>1.45 [0.90, 2.03]</td>
</tr>
<tr>
<td>+ Seo 2014</td>
<td>96</td>
<td>1192</td>
<td>1.39 [0.91, 2.13]</td>
</tr>
<tr>
<td>+ Guo 2015</td>
<td>98</td>
<td>1277</td>
<td>1.35 [0.90, 2.03]</td>
</tr>
<tr>
<td>+ Kogure 2015</td>
<td>102</td>
<td>1359</td>
<td>1.35 [0.90, 2.03]</td>
</tr>
<tr>
<td>+ Omar 2017</td>
<td>105</td>
<td>1420</td>
<td>1.35 [0.90, 2.03]</td>
</tr>
</tbody>
</table>

*Cumulative number of patients are presented

---

**Fig. 7** Forest plots showing the results of cumulative meta-analyses of endoscopic papillary balloon dilation (EPBD) vs. endoscopic sphincterotomy (EST) for the removal of common bile duct stones. a Post-endoscopic retrograde cholangiopancreatitis pancreatitis (PEP). b Overall short-term complications. In the cumulative meta-analysis, a summary odds ratio (OR) was calculated each time the results of a new study were added. Although significance was not achieved, PEP was more common after EPBD than after endoscopic sphincterotomy since 2003, whereas the incidence of overall short-term complications did not differ over time between the two endoscopic techniques. CI, confidence interval.
Some authors [52–53] have suggested that endoscopic sphincterotomy could lead to long-term complications such as cholecystitis and recurrence of CBD stones because of damage to papillary function. Based on the results of our meta-analysis, there was a trend, without statistical significance, toward an increased risk of long-term cholangitis, cholecystitis, and stone recurrence in patients undergoing endoscopic sphincterotomy compared with those undergoing EPBD. Moreover, EPBD may be preferred in terms of patient comfort, as it requires fewer sessions of repeat ERCP compared with endoscopic sphincterotomy. This in turn would result in cost savings for the healthcare system, as fewer reinterventions are needed [54].

This meta-analysis has some limitations. Our literature search considered several medical literature databases allowing us to search across conference proceedings, but we chose to consider only English language publications for inclusion in this systematic review and meta-analysis. It is unlikely, however, that the potential exclusion of non-English publications has biased our results. In fact, previous investigations generally found no evidence of bias from the use of language restrictions in systematic review-based meta-analyses [55–56]. We did not try to contact study investigators to retrieve additional data because a previous study showed unsuccessful attempts [57]. From a statistical perspective, we used the OR instead of the relative risk as an association measure to enable comparison between our results and those from previously published meta-analyses [4–7, 10–13]. Between-study heterogeneity may be a concern in meta-analyses [20]. Thus, we performed stratified analyses according to the geographic area of the study, stone size, balloon diameter, and dilatation time to explore potential sources of heterogeneity. Indeed, the limited number of studies reporting data according to stone size represents a limitation, while the use of a priori data-driven cutoffs for balloon diameter (10 mm) and balloon dilatation time (1 minute) may have introduced bias.

Our results highlight the need of further RCTs aimed at studying efficacy and safety of EPBD vs. endoscopic sphincterotomy by balloon diameter, dilatation time, stone size, number of attempts, cannulation time, and stone recurrence rate. Moreover, new studies should perform subgroup analyses that also consider patients with altered anatomy and peripancreatic diverticulum.

In conclusion, although our latest data confirm that EPBD is currently inferior to endoscopic sphincterotomy in terms of overall stone clearance, this is the first meta-analysis to report an absence of evidence of a difference in efficacy when EPBD was performed with large (≥10 mm) balloons. EPBD with large balloons (≥10 mm) also increased clearance and reduced the...

![Fig. 9](thumb) Forest plots showing the results from subgroup analyses according to the geographic area of the study (Asian vs. Western), balloon diameter (<10 vs. ≥10 mm), balloon dilatation time (<1 vs. ≥1 minute), and stone size (<10 vs. ≥10 mm). a Overall stone clearance. b Use of mechanical lithotripsy. c Post-endoscopic retrograde cholangiopancreatitis.
need for EML, while longer dilation time (>1 minute) reduced pancreatitis risk. For these reasons, EPBD should be considered as an alternative treatment for CBD stone removal, especially in patients with altered anatomy, perivaterian diverticulum, coagulopathy, and cholecystitis, and in patients with higher risk of bleeding.

Acknowledgment
The work of Marta Rossi was supported by the Italian Association for Cancer Research (AIRC) (My First AIRC Grant No. 17070).

Competing interests
None.

References


[46] Youn YH, Lim HC, Jahng JH et al. The increase in balloon size to over 15 mm does not affect the development of pancreatitis after endoscopic papillary large balloon dilatation for bile duct stone removal. Dig Dis Sci 2011; 56: 1572 – 1577


[48] Liao WC, Lee CT, Chang CY et al. Randomized trial of 1-minute versus 5-minute endoscopic balloon dilation for extraction of bile duct stones. Gastrointest Endosc 2010; 72: 1154 – 1162


