Is Endoscopic Balloon Dilation Still Associated With Higher Rates of Pancreatitis?

A Systematic Review and Meta-Analysis

Carolina Ogawa Matsubayashi, MD, Igor Braga Ribeiro, MD,* Diogo Turiani Hourneaux de Moura, MD, MSc, PhD,*† Vitor Ottoboni Brunaldi, MD,* Wanderley Marques Bernardo, MD, PhD,* Kelly E. Hathorn, MD,† and Eduardo Guimarães Hourneaux de Moura, PhD**

Abstract: The objective of this study was to compare the efficacy and safety of endoscopic papillary balloon dilation (EPBD), endoscopic sphincterotomy (ES), and the combination of large balloon dilation and ES (ES + EPLBD) in the treatment of common bile duct stones, with a special focus on postendoscopic retrograde cholangiopancreatography (ERCP) pancreatitis (PEP). Individualized search strategies were developed in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. We included randomized controlled trials (RCTs) which evaluated at least one of the following outcomes: PEP, complete stone removal in the first ERCP, need for mechanical lithotripsy, recurrence of common bile duct stones, bleeding, and cholangitis. Twenty-five RCTs were selected for analysis. Pancreatitis rates were higher for EPBD than for ES ($P = 0.003$), as were severe pancreatitis rates ($P = 0.04$). However, in the 10-mm or greater balloon subgroup analysis, this difference was not shown ($P = 0.82$). Rates of PEP were higher in the subgroup of non-Asian subjects ($P = 0.02$), and the results were not robust when RCTs that used endoscopic nasobiliary drainage were omitted. The incidence of pancreatitis was comparable between EPLBD and ES + EPLBD. All 3 approaches were equally efficacious. Nevertheless, the results should be interpreted with caution, because pancreatitis is a multifactorial pathology, and RCTs can have limited generalizability.

Key Words: common bile duct stone, balloon dilation, endoscopic papillary balloon dilatation; endoscopic sphincterotomy, lithotripsy

(*Pancreas* 2020;49: 158–174)

Pancreatitis is one of the major complications in endoscopic retrograde cholangiopancreatography (ERCP), with an incidence of 3% to 15%,[1] depending on the risk factors related to the patient and the procedure. One such factor is the type of access to the biliary tract for removal of stones from the common bile duct (CBD), which can be achieved through the use a various techniques, the standard technique being endoscopic sphincterotomy (ES).[2] In complex cases, it can be necessary to use a combination of methods.[3,4]

Endoscopic papillary balloon dilation (EPBD), first described by Staritz et al,[5] was introduced as an alternative to ES for the removal of bile duct stones, to minimize the adverse effects of the procedure. However, the method has become obsolete, especially in non-Asian countries, due to the higher incidence of post-ERCP pancreatitis (PEP) after EPBD.[6–8] More recently, EPBD has been indicated mainly for patients with coagulopathy and a history of gastrointestinal surgery in whom ES is risky or technically challenging.[9] However, ES alone is less efficacious for the treatment of large bile duct stones and mechanical lithotripsy (ML) is often needed. In this context, Ersoz et al[10] modified the technique by using large balloon EPBD after ES (ES + EPLBD), achieving good results. Although subsequent studies supported those findings, the best choice of technique and type of access remain controversial.[11,12] Recently, some authors have adopted EPLBD without precutting the papilla, reporting good rates of bile duct clearance under safe conditions, with no increased risk of pancreatitis.[13–15] Consequently, the real incidence of PEP due to balloon dilation itself has come into question.[16]

Although the management of EPLBD is explained in current guidelines, some aspects remain unclear, especially regarding pancreatitis rates.[17,18] In addition, since the last meta-analysis,[15] more trials comparing EPLBD with other techniques have been published. In this context, the primary purpose of this systematic review and meta-analysis was to compare EPLBD, ES, and ES + EPLBD for the treatment of choledocholithiasis, exploring the incidences rates for PEP. A secondary aim was to evaluate success rates, the need for ML, and other post-ERCP complications.

METHODOLOGY

This systematic review and meta-analysis was conducted in accordance with recommendations outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.[19] It was also registered in the International Prospective Register of Systematic Reviews database (Registration no. CRD42018111392).

Inclusion Criteria

We included only randomized controlled trials (RCTs) that met all of the patient-intervention-comparison-outcome criteria, defined as follows: patient (patients who underwent ERCP for the treatment of CBD stones); intervention (EPBD or EPLBD); comparison (ES or ES + EPLBD); and outcome (evaluating at least 1 outcome). The primary outcome was PEP. The secondary outcomes were complete stone removal at the first ERCP, the need for ML, recurrence of bile duct stones, and other post-ERCP complications (bleeding and cholangitis). Only full-text articles were considered eligible.

From the *Gastrointestinal Endoscopy Unit, Department of Gastroenterology, University of São Paulo School of Medicine, São Paulo, Brazil; and †Division of Gastroenterology, Hepatology and Endoscopy Brigham and Women's Hospital, Harvard Medical School, Boston, MA.
Received for publication July 27, 2019; accepted December 18, 2019.
Address correspondence to: Igor Braga Ribeiro, MD, Academic Research, Gastrointestinal Endoscopy Unit, Department of Gastroenterology, University of São Paulo School of Medicine, R. Dr. Ovídio Pires de Campos, 255-Cerqueira César, São Paulo 05403-010, Brazil (e-mail: igorbraga1@gmail.com).
The authors declare no conflict of interest.

DOI: 10.1097/MPA.0000000000001489

Search Strategy

We conducted a comprehensive, systematic search of the scientific literature to identify all published and unpublished RCTs. To that end, we searched the Excerpta Medica, MEDLINE (PubMed), and Cochrane Central Register of Controlled Trials databases for RCTs published up through June 2019. We used the following search strings: (*calcul** OR *stone** OR *lithiasis* OR *gallstone* OR *choledocholithiasis* OR *common bile duct diseases* OR *cholelithiasis*) AND (*endoscopic papillary balloon dilatation* OR *balloon dilation* OR *balloon catheter* OR *papillotomy* OR *Sphincteroplasty* OR *Sphincteroplasties* OR *sphincterotomy* OR *sphincterotomies* OR *precut* OR *Papillotomies*).

Study Selection and Data Extraction

Two authors, working independently, reviewed all titles and abstracts for potential inclusion, as well as the full texts of the articles selected. A third author resolved any disagreements. The following data were retrieved from each article: first author; publication year; number of patients; baseline characteristics; risk of bias; interventions; and outcomes, which were collected on an intention-to-treat basis when possible.

Our study involved 2 different groups: Balloon dilation versus ES (EPBD or EPLBD compared with ES); and EPLBD versus EPLBD + ES. We accepted the author definitions of PEP and extracted the overall incidence of pancreatitis. In addition, we ran a separate analysis of the severe cases, classified preferably according to Cotton et al.[20] Bleeding was recorded only in cases of clinical (not just endoscopic) evidence of bleeding, or when it occurred 48 hours after the procedure, or when there was a ≥2 g/dL decrease in hemoglobin levels, or when there was a need for blood transfusion.

Assessment of Risk of Biases and Quality of Evidence

Eligible RCTs were assessed for biases using a standardized table regarding randomization, allocation, losses, outcomes, selective reporting, sample size calculation, and intention-to-treat analysis. We assessed the overall quality of the evidence for each outcome, grading it as very low, low, moderate, or high, according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria, using the GRADEpro Guideline Development Tool.[21] In judging the quality of the evidence, we examined the following variables: risk of bias; inconsistency of results; imprecision; indirectness; publication bias; dose–effect response; magnitude of the effect; and plausible confounders.

Data Analysis

We performed statistical analyses using Review Manager, version 5 (RevMan 5, The Cochrane Collaboration, Copenhagen, Denmark, 2014). Pooled estimates were calculated with the fixed effect model devised by Mantel and Haenszel[22] and presented with 95% confidence intervals (CIs). Risk differences (RDs) were used to compare all outcomes. Heterogeneity was assessed using the Higgins test (I^2),[23] and we considered the I^2 interpretation thresholds suggested in the Cochrane Handbook.[24] For outcomes with high heterogeneity, we initially inspected the forest plots and funnel plots to identify publication bias. If an outlier study was not detected and bias could not be attributed to another factor, true heterogeneity was presumed and the pooled estimates were computed using the random-effects model devised by DerSimonian and Laird.[25]

For the main outcome, a leave-one-out analysis was carried out by removing 1 study at a time to examine the influence that each individual study had on our estimate.[26] For all outcomes, we performed sensitivity analysis by excluding studies with

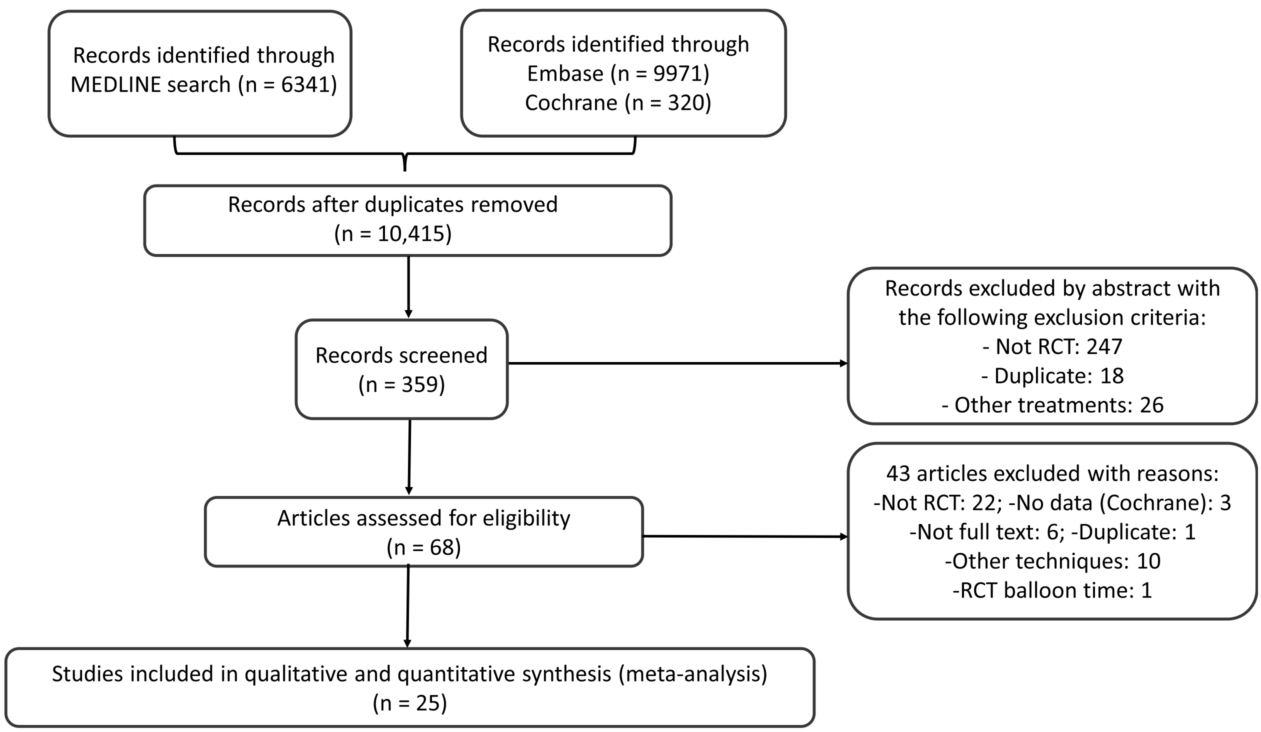

FIGURE 1. Study flow diagram.

TABLE 1A. Balloon Dilation Versus Sphincterotomy: Baseline Characteristics of the Patients

Author	Sample Size Total (EPBD/ES), N (n/n)	Female (EPBD/ES) %/%	PAD (EPBD/ES) %/%	Age, y*		No. Stones*		Stone Diameter, mm*		CBD Caliber, mm*	
				EPBD	EPBD + ES	EPBD	EPBD + ES	EPBD	EPBD + ES	EPBD	EPBD + ES
Arnold et al, 2001[45]	60 (30/30)	63.3/56.7	NR	54.2 (SD, 18.5) (23–85)	58.5 (SD, 18.5) (22–94)	1.6 (SD, 1.1)	1.8 (SD, 1.5)	7 (SD, 3.5) (3–15)	10 (SD, 4.7) (5–20)	NR	NR
Bergman et al, 2001[48]	34 (16/18)	25.0/0.0	6.0/0.0	73 (43–84)	72 (61–84)	2 (1–10)	2 (1–10)	9 (5–22)	8 (4–20)	NR	NR
Bergman et al, 2001[49]	180 (93/87)	61.3/64.4	23.0/16.0	67 (SD, 16.4)	67 (SD, 16.0)	1 (1–10)	1 (1–10)	9 (3–37)	8 (3–27)	NR	NR
Bergman et al, 1997[57]	202 (101/101)	57.4/55.4	22.0/15.0	72 (27–98)	71 (29–96)	2 (1–14)	1 (1–15)	10 (3–36)	9 (4–27)	NR	NR
Chu et al, 2017[50]	62 (30/33)	56.7/54.5	46.7/34.4	64.7 (SD, 6.5)	65.6 (SD, 7.4)	≥3: 19 (63.3)	≥3: 18 (56.2)	<15: 21 (64) ≥15: 9 (36)	<15: 24 (75) ≥15: 8 (25)	18.4 (SD, 5.8)	17.9 (SD, 5.5)
Disario et al, 2004[6]	237 (117/120)	65.0/74.2	NR	47 (SD, 19)	54 (SD, 19)	1 (1–100)	1 (1–10)	6 (0.5–10)	5 (0.5–14)	10 (4–20)	10 (1–25)
Fu et al, 2013[52]	206 (103/103)	49.5/56.3	NR	61.8 (SD, 17.4)	60.5 (SD, 14.7)	2.17 (SD, 1.43)	1.89 (SD, 1.37)	8.38 (SD, 2.67)	7.71 (SD, 2.35)	12.74 (SD, 2.79)	12.55 (SD, 3.05)
Fujita et al, 2003[51]	282 (138/144)	45.7/36.1	42/45	66.8 (26–93)	68.4 (31–93)	2.4 (SD, 2.9)	2.4 (SD, 2.5)	7.0 (SD, 3.1)	7.3 (SD, 3.4)	11.8 (SD, 3.8)	12.7 (SD, 4.1)
Guo et al, 2015[53]	170 (85/85)	47.1/49.4	31.0/32.0	62 (SD, 17)	59 (SD, 16)	>3: 21 (25)	>3: 23 (27)	10 (10–30)	10 (10–40)	12 (11–30)	12 (11–40)
Kogure et al, 2015[34]	171 (86/85)	55.8/51.8	56.0/67.0	79.1 (SD, 8.1)	80.9 (SD, 8.2)	2–3: 27 (31) ≥4: 25 (29)	2–3: 29 (34) ≥4: 29 (34)	15.2 (SD, 4.6)	14.3 (SD, 4.8)	16.4 (SD, 3.6)	16.0 (SD, 4.1)
Lin et al, 2004[46]	104 (51/53)	45.1/41.5	41.0/38.0	64 (28–90)	65 (28–88)	≤3: 31 (61) >3: 20 (39)	≤3: 25 (47) >3: 28 (53)	8 (SD, 6)	8 (SD, 6)	NR	NR
Minakari et al, 2013[39]	160 (80/80)	51.3/47.5	NR	56.4 (SD, 15.3)		NR	NR	NR	NR	NR	NR
Minami et al, 1995[54]	40 (20/20)	35.0/55.0	NR	64.0 (SD, 11.2) (38–82)	71.3 (SD, 14) (44–83)	NR	NR	NR	NR	NR	NR
Natsui et al, 2002[35]	140 (70/70)	52.9/52.9	54.0/61.0	64.5 (23–87)	67.1 (38–88)	2.7 (1–15)	2.6 (1–15)	9.2 (3–22)	9.7 (3–17)	NR	NR
Ochi et al, 1999[36]	110 (55/55)	38.2/43.6	34.5/25.5	62.6 (SD, 15.9) (24–87)	66.3 (SD, 14.3) (21–89)	2.1 (SD, 9)	1.7 (SD, 1.2)	8.1 (SD, 3.4)	8.8 (SD, 4.2)	NR	NR
Oh and Kim, 2012[38]	83 (40/43)	50.0/46.5	57.5/40.0	72.3 (SD, 9.5)	68.7 (SD, 12.9)	1: 20 (50)	1: 28 (65)	13.2 (SD, 3.6)	13.1 (SD, 3.9)	18.0 (SD, 4.3)	18.2 (SD, 4.6)
Omar et al, 2017[37]	124 (61/63)	57.4/60.3	14.3/21.3	47.8 (SD, 14.5)	44.8 (SD, 13.9)	2.3 (SD, 1.5)	2.1 (SD, 1.4)	13.9 (SD, 2.4)	13.1 (SD, 2.6)	18.2 (SD, 4.6)	17.1 (SD, 4.3)
Seo et al, 2013[47]	132 (62/70)	56.5/54.3	NR	32.1 (SD, 7.3)	33.2 (SD, 5.8)	1.5 (1–5)	1.8 (1–8)	7.2 (SD, 2.08)	7.6 (SD, 3.12)	NR	NR
Takezawa et al, 2004[43]	91 (46/45)	30.4/33.3	34.0/42.0	70 (40–90)	69 (41–93)	1 (1–7)	1 (1–7)	10 (1–35)	11 (3–27)	14 (9–28)	16 (8–32)
Tanaka et al, 2004[44]	32 (16/16)	37.5/18.8	50.0/37.5	67.2 (50–78)	70.6 (49–87)	2 (1–12)	2 (1–4)	10.2 (SD, 3.5) (5–15)	12.4 (SD, 6) (4–24)	NR	NR
Vlavianos et al, 2003[41]	202 (103/99)	75.7/64.6	NR	60.8 (56.8–64.7)	61.9 (58.3–65.4)	>3: 40 (38.8)	>3: 25 (25.3)	<5: 27 6–9: 37 >10: 39	<5: 31 6–9: 26 >10: 42	Dilated to >7: 83	Dilated to >7: 77
Watanabe et al, 2007[42]	180 (90/90)	43.3/45.6	42.0/39.0	69.1 (SD, 13.1) (26–96)	70.2 (SD, 8.1) (34–88)	2.7 (SD, 2.8) (1–19)	2.5 (SD, 2.7) (1–14)	8.1 (SD, 3.2) (2–20)	7.7 (SD, 2.9) (2–20)	NR	NR
Yasuda et al, 2001[40]	70 (35/35)	54.3/40.0	NR	69.5 (42–86)	69.4 (43–88)	3.7 (1–16)	3.3 (1–16)	12.4 (4–24)	12.3 (5–24)	15.1 (6–30)	14.7 (6–30)

TABLE 1B. Balloon Dilation Versus Balloon Dilation Plus Sphincterotomy: Baseline Characteristics of the Patients

Author	Sample Size Total (EPBD/ES), N (n/n)	Female (EPBD/ES) %/%	PAD (EPBD/ES) %/%	Age, y*		No. Stones*		Stone Diameter, mm*		CBD Caliber, mm*	
				EPBD	EPBD + ES	EPBD	EPBD + ES	EPBD	EPBD + ES	EPBD	EPBD + ES
Park et al, 2018[27]	200 (100/100)	55.0/52.0	NR	74 (56–91)	73 (49–91)	2 (1–19)	2 (1–28)	15.2 (14.1–16.2)	14.6 (13.8–15.5)	17.8 (11–29.7)	18 (10–45)
Cheon et al, 2017[55]	86 (42/44)	50.0/45.5	50.0/65.9	71.0 (SD, 12.4)	71.7 (SD, 10.1)	2.6 (6.9)	2.6 (2.3)	14.4 (SD, 3.3)	14.0 (SD, 2.1)	15.8 (3.6)	16.1 (3.2)
Guo et al, 2015[56]	170 (85/85)	47.1/45.9	30.6/54.1	62 (SD, 17)	63 (SD, 16)	>3: 21 (25)	>3: 24 (28)	10 (10–30)	10 (10–30)	12 (11–30)	12 (11–30)
Chu et al, 2017[50]	66 (30/33)	56.7/54.5	46.7/36.4	64.7 (SD, 6.5)	64.8 (SD, 5.5)	≥3: 19 (63)	≥3: 20 (61)	<15: 21 (64) ≥15: 9 (36)	<15: 19 ≥15: 14	18.4 (SD, 5.8)	18.1 (SD, 4.2)

*Results presented as mean (SD) (range), mean (SD), median (range), [value]: n (%), or [value]: n.

PAD indicates periampullary diverticula; NR, not reported; SD, standard deviation.

unclear or high selection bias. We also investigated whether the rates of PEP and severe PEP would be different if patients received only endoscopic nasobiliary drainage (ENBD) as a factor of protection.

Given the difference between interventions and based on previous statements, we prespecified subgroup analyses for the outcomes, pancreatitis, need for ML, and stone removal, by balloon size (≥10 mm vs <10 mm). Regarding PEP, we also performed a subgroup analysis separating the RCTs conducted in non-Asian countries and those conducted in Asian countries, because it has been suggested that ethnic differences can have an influence.[27]

RESULTS

The study selection process yielded 10,415 studies (Fig. 1). Initially, 10,391 studies were excluded. Of those, 7 were excluded because the full text was not available.[28–34] However, for one of those 7 studies, which was published only as a conference abstract, we were able to obtain the full text directly from the author.[34] Therefore, the meta-analysis included 25 full-text RCTs.

The balloon dilation versus ES group comprised 22 studies,[6,34–54] and the EPLBD versus ES + EPLBD group comprised 4 studies,[27,50,53,55] with collective totals of 2871 and 519 subjects, respectively. Two studies performed both comparisons and were therefore included in both groups.[50,53] The characteristics of the 25 articles are described in Tables 1 and 2.

Because of the nature of the interventions, none of the endoscopists performing the procedure were blinded to the arm of the study to which a given patient had been assigned. There were 3 studies with substantial (>20%) losses to follow-up.[37,41,43] In 3 studies, there was a baseline characteristic that differed significantly between the balloon arm and the comparison arm and could have led to some degree of bias: younger patients in the balloon dilation arm[6,50]; stones greater than 15 mm in the balloon dilation arm; and a higher proportion of patients with cirrhosis in the balloon dilation arm.[42] In another 2 studies,[39,54] the authors did not include data for the size/number of stones or the diameter of the CBD, so we considered the distribution of the patients unclear. For all 25 studies, quality and risk of bias are described in detail (Fig. 2).

Comparison 1: Balloon Dilation Versus ES

Pancreatitis

Of the 22 studies in the balloon dilation versus ES group, only one was not eligible for inclusion in the calculation of the effect sizes for pancreatitis and severe pancreatitis (Figs. 3 and 4).[43] Four studies[34,38,45,50] defined severity in a manner different from that outlined by Cotton et al,[20] and 3 others did not specify the criteria applied.[40,46,54]

In the initial analysis, balloon dilation significantly increased the incidence of PEP in comparison with ES (RD, 0.03; 95% CI, 0.01–0.04; $P = 0.003$). An I^2 of 51% was considered moderate because, although the direction of the effect is clear, the confidence interval is narrow. Severe cases were also significantly more common in the balloon dilation arms (RD, 0.01; 95% CI, 0.00–0.02; $P = 0.04$; I^2, 0%).

Sensitivity Analysis

As shown in Table 3, the leave-one-out sensitivity analysis showed that after exclusion of the Disario et al study,[6] pancreatitis rates had a tendency to be higher in the balloon dilation arms, although there was no statistical significance (RD, 0.02; 95% CI, −0.00 to 0.03; $P = 0.10$; I^2, 25%). For the incidence of PEP and severe PEP, the results were robust after 6 studies with selection

TABLE 2A. Balloon Dilation Versus Sphincterotomy: Characteristics of the Procedures

Author	Balloon Size, mm	Use of ENBD	Use of NSAIDs	Other Characteristics	Pancreatitis Score Classification
Arnold et al, 2001[45]	8		NR		WHO
Bergman et al, 2001[48]	8		NR	Use of needle knife	Cotton et al[20]
Bergman et al, 2001[49]	8		No	Pancreatography: EPBD = 64 (69%); ES = 54 (62%)	Cotton et al[20]
Bergman et al, 1997[57]	8		No	When complete stone removal was not possible after EBD or after a maximum of 60 min, an additional sphincterotomy was done as an "escape" procedure	Cotton et al[20]
Chu et al, 2017[50]	Range, 12–20		In all patients	Prophylactic pancreatic stents used on a case-by-case basis	Freeman et al[7]
Disario et al, 2004[6]	8		NR	Pancreatography: EPBD = 57 (49%); ES = 57 (48%)	Cotton et al[20]
Fu et al, 2013[52]	Range, 10–12	Post-procedure, in all patients	NR	Did not exclude previous ES, use of precut ES	Cotton et al[20]
Fujita et al, 2003[51]	Range, 4–8		Post-procedure gabexate: EPBD = 137 (99%); ES = 137 (95%)	ES after precut ES, when deep cannulation of the bile duct was not possible	Cotton et al[20]
Guo et al, 2015[53]	Range, 10–15	Post-procedure, in all patients	NR		Cotton et al[20]
Kogure et al, 2015[34]	Range, 15–18		NR		ASGE lexicon
Lin et al, 2004[46]	Range, 8–12		NR		Unclear
Minakari et al, 2013[39]	Range, 12–15		NR		Cotton et al[20]
Minami et al, 1995[54]	Not clear	EPBD = 3 patients	NR		Unclear
Natsui et al, 2002[35]	8	After incomplete stone removal	NR		Cotton et al[20]
Ochi et al, 1999[36]	8		No		Cotton et al[20]
Oh and Kim, 2012[38]	Range, 10–18	After incomplete stone removal	Gabexate mesylate in all patients	Pancreatography: EPBD = 5 (13%); ES = 8 (19%)	Ranson+Glasgow
Omar et al, 2017[37]	Range, 12–15	After incomplete stone removal: EPBD = 6 (61%); ES = 14 (22%)	NR		Cotton et al[20]
Seo et al, 2013[47]	Range, 6–10		NR		Cotton et al[20]
Takezawa et al, 2004[43]	8		NR		Cotton et al[20]
Tanaka et al, 2004[44]	8	Indicated for pancreatitis and cholangitis: EPBD = 1; ES = 6	Indicated for pancreatitis and cholangitis: EPBD = 1; ES = 6		Cotton et al[20]
Vlavianos et al, 2003[41]	10		NR	Precut ES if cannulation was not possible	Cotton et al[20]
Watanabe et al, 2007[42]	8		Intravenous gabexate mesylate for 3 days	Liver cirrhosis: EPBD = 5	Cotton et al[20]
Yasuda et al, 2001[40]	10				Unclear

WHO indicates World Health Organization; ASGE, American Society for Gastrointestinal Endoscopy.

bias had been excluded.[36,44–46,50,54] However, after 6 RCTs in which patients benefited from ENBD had been omitted,[35,37,38,52–54] the rates of PEP and severe PEP were significantly higher in the balloon dilation arms.

Subgroup Analysis

The subgroup analysis of balloon size showed significantly higher rates of pancreatitis in the <10 mm balloon subgroup (RD, 0.06; 95% CI, 0.03–0.08; $P < 0.0001$; I^2, 56%; Fig. 5). In the subgroup of studies conducted in non-Asian countries,[6,37,41,45,48,49] the rates of pancreatitis were also significantly higher (RD, 0.06; 95% CI, 0.02–0.09; $P = 0.0009$; I^2, 61%; Fig. 6). For severe pancreatitis (Figs. 7 and 8), the non-Asian country subgroup was the only subgroup that demonstrated a statistical difference (RD, 0.02; 95% CI, 0.00–0.04; $P = 0.03$; I^2, 0%), including 9 cases in balloon dilation arms and one in an ES arm.

Stone Clearance in a Single Session

There were only 3 studies that provided no information regarding stone clearance.[39,43,54] Stones were removed at first

TABLE 2B. Balloon Dilation Versus Balloon Dilation Plus Sphincterotomy: Characteristics of the Procedures

Author	Balloon Size, mm	Use of ENBD	Use of NSAIDs	Other Characteristics	Pancreatitis Score Classification
Park et al, 2018[27]	10–20		No		Cotton et al[20]
Cheon et al, 2017[55]	12–14		NR		Cotton et al[20]
Guo et al, 2015[53]	10–15	Post-procedure, in all patients	NR		Cotton et al[20]
Chu et al, 2017[50]	12–20		In all patients	Prophylactic pancreatic stents used on a case-by-case basis	Freeman et al[7]

ERCP in 1053 of 1281 subjects when balloon dilation of the papillae was performed and in 81.5% when ES was performed, the success rate being similar in both groups (RD, 0.01; 95% CI, −0.02 to 0.04; P = 0.59; I^2, 70%;). We did not detect any outliers, and we therefore calculated the RD using a random-effects model (RD, −0.01; 95% CI, −0.05 to 0.04; P = 0.80; Fig. 9).

Subgroup Analysis

The differences among the balloon size subgroups were significant (P = 0.07). Among the 10 RCTs evaluated in the <10 mm balloon subgroup,[6,35,36,40,42,44,45,48,49,51] there was a tendency toward higher success rates for ES (RD, 0.06; 95% CI, −0.02 to 0.13; P = 0.14; I^2, 68%; Fig. 10). In contrast, among the 8 RCTs evaluated in the ≥10 mm balloon subgroup,[34,37,38,41,50,52,53] there was a tendency toward higher success rates for balloon dilation (RD, −0.03; 95% CI, −0.10 to 0.03; P = 0.29; I^2, 63%). The quality of evidence was downgraded by 1 level due to high unexplained inconsistency.

ML

Although the need for ML was more common in the balloon dilation arms, the difference in comparison with the ES arms was not significant (RD, −0.02; 95% CI, −0.00 to 0.05; P = 0.06; I^2, 61%; Fig. 11). Because no publication bias was detected, we incorporated heterogeneity into random-effects models (RD, 0.02; 95% CI, −0.01 to 0.05; P = 0.24; I^2, 61%).

Subgroup Analysis

The results of the test for subgroup differences suggested that there is a statistically significant subgroup effect regarding balloon size (P = 0.03; Fig. 12). The need for ML was more common in the <10 mm balloon subgroup (RD, 0.09; 95% CI, 0.01–0.18; P = 0.03). However, there was substantial heterogeneity between the RCTs. Therefore, the validity of this effect estimate is uncertain, as trial results are inconsistent.

In the >10 mm balloon subgroup, there was more use of ML in the ES arms, although the difference was not significant (RD, −0.02; 95% CI, −0.07 to 0.03; P = 0.43). However, we did not downgrade it for imprecision because the confidence interval did not fail to exclude the potential benefit.

Bleeding

Twenty-one studies with a collective total of 2669 patients were included in the bleeding analysis.[34,35,37–42,44–54] The incidence of bleeding was significantly higher in the ES arm (RD, −0.02; 95% CI, −0.03 to −0.01; P = 0.002; Fig. 13), with a number needed to harm of 50. The results of this analysis were extremely homogeneous among the RCTs (I^2, 0%; P = 0.57). Therefore, we judged this evidence to be of high quality.

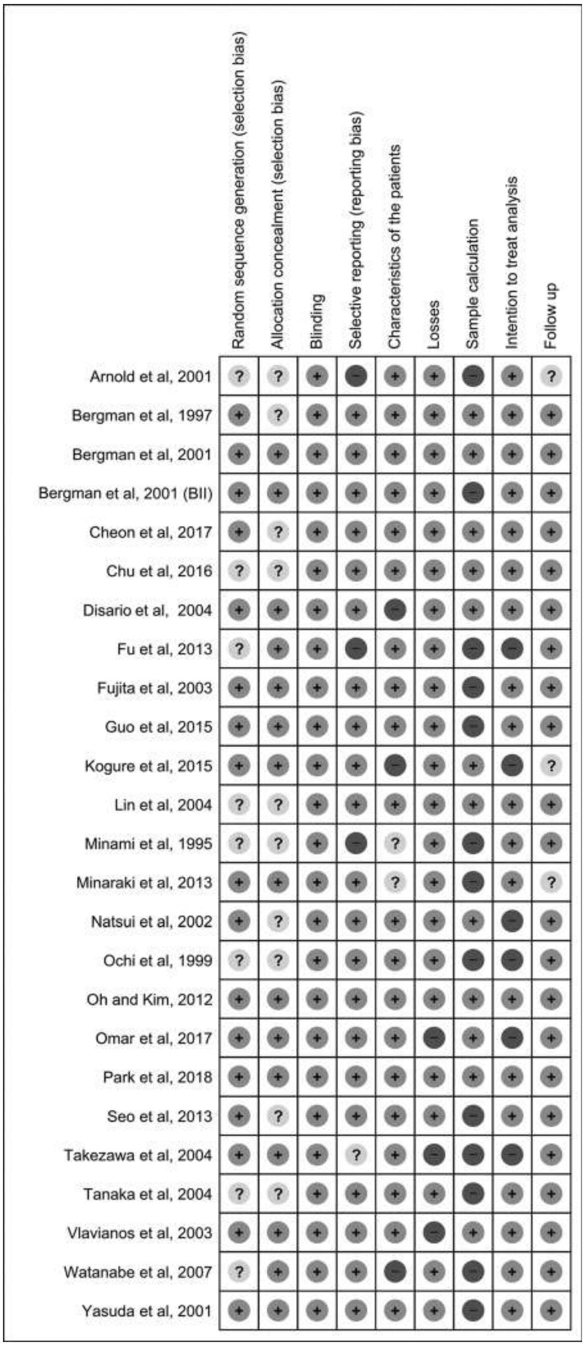

FIGURE 2. Risk of bias summary: review author judgments about each risk of bias for each study included.

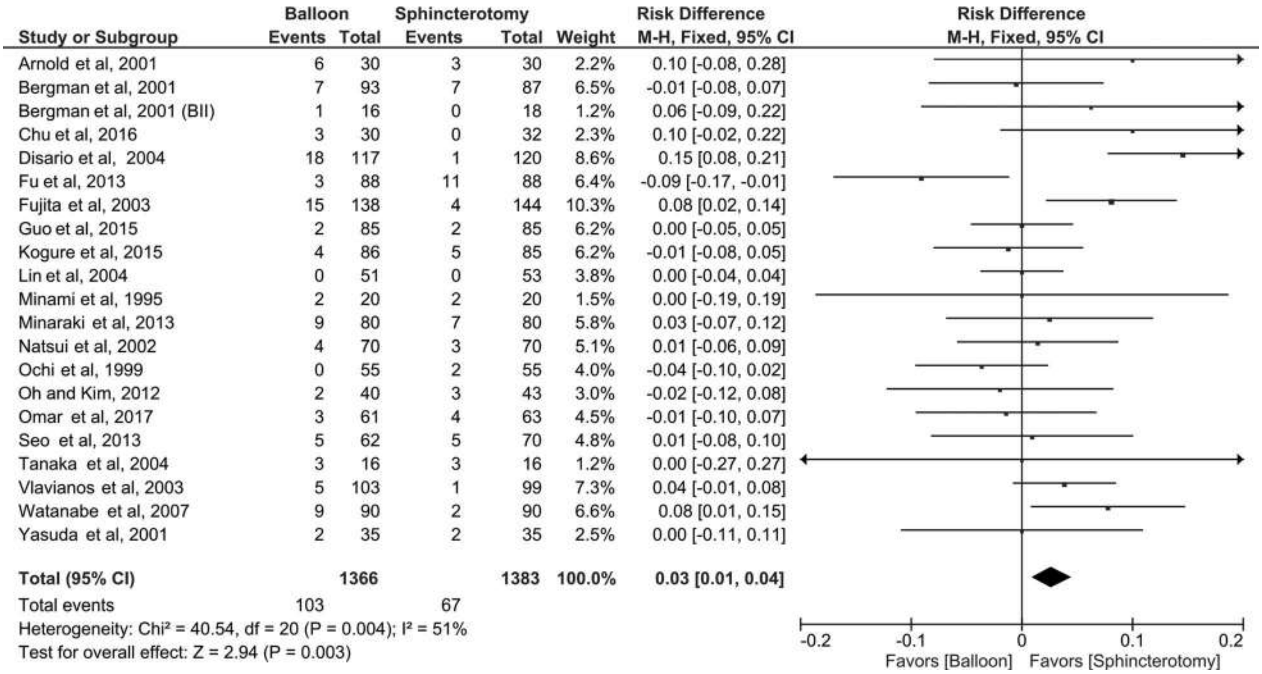

FIGURE 3. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of pancreatitis. df, degrees of freedom; M-H, Mantel-Haenszel.

Cholangitis

A total of 1938 patients from 15 studies were evaluated for cholangitis,[6,34–38,40–45,50,52,53] and there was no significance difference between arms (RD, 0.01; 95% CI, −0.01 to 0.02; $P = 0.25$; I^2, 0%; Fig. 14).

Recurrence

Recurrence was analyzed in 12 studies,[6,35–37,40,41,43,44,46,47,50,57] with collective totals of 766 and 778 patients in the balloon dilation and ES arms, respectively. Forty events were reported in each group, and there was therefore no difference between the methods (RD, −0.00; 95% CI, −0.02 to −0.02; $P = 0.95$; I^2, 0%; $P = 0.68$; Fig. 15).

Sensitivity Analyses

For all secondary outcomes, the results of the sensitivity analyses were robust after studies with a selection bias had been excluded (Table 3).

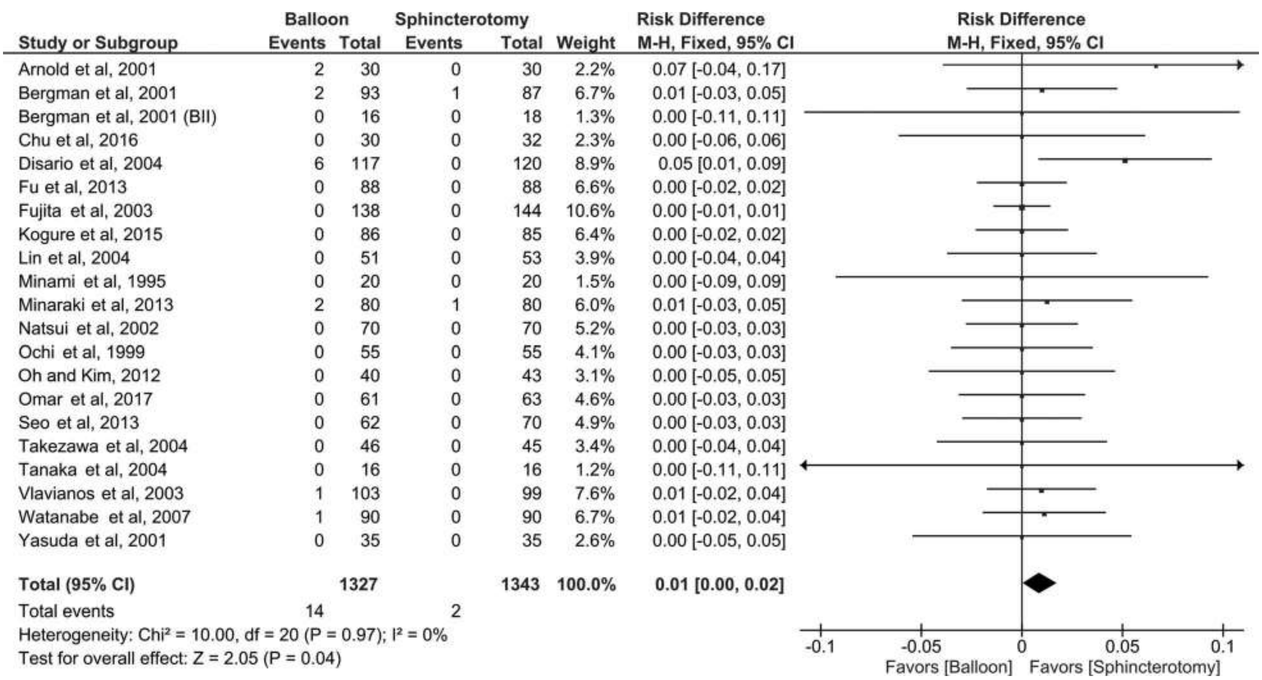

FIGURE 4. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of severe pancreatitis.

TABLE 3. Sensitivity Analysis

Analysis	Preremoval				Postremoval		
Outcome	**n**	**RD (95% CI)**	**I^2, %**	**Studies Omitted***	**n**	**RD (95% CI)**	**I^2, %**
Leave-one-out							
Pancreatitis							
EPBD vs ES	21	0.03 (0.01–0.04)	51	6	20	0.02 (−0.00 to 0.03)	25
EPLBD vs ES + EPLBD EPBD	4	−0.00 (−0.04 to 0.03)	26	50	3	−0.02 (−0.05 to 0.02)	0
Selection bias				36,44,45,46,50,54			
Pancreatitis	21	0.03 (0.01–0.04)	51		15	0.03 (0.01–0.05)	56
Severe pancreatitis	21	0.00 (−0.00 to 0.01)	0		15	0.00 (−0.00 to 0.01)	0
Lithotripsy	19	0.02 (−0.01 to 0.05)	61		12	0.01 (−0.02 to 0.05)	63
EPBD vs ES							
Removal in 1 ERCP	19	−0.01 (−0.05 to 0.04)	70		14	0.02 (−0.02 to 0.06)	65
Bleeding	20	−0.02 (−0.03 to −0.01)	0		15	−0.02 (−0.03 to −0.01)	0
Recurrence	12	0.00 (−0.02 to 0.02)	0		8	0.01 (−0.02 to 0.03)	0
Cholangitis	15	0.01 (−0.01 to 0.02)	0		11	0.01 (−0.01 to 0.02)	0
Use of ENBD				35,37,38,52,53,54			
Pancreatitis							
EPBD vs ES	21	0.03 (0.01–0.04)	51		15	0.05 (0.02–0.07)	55
EPLBD vs ES + EPLBD	4	−0.00 (−0.04 to 0.03)	26		3	−0.00 (−0.05 to 0.04)	50
Severe pancreatitis							
EPLBD vs ES + EPLBD	21	0.00 (−0.00 to 0.01)	0		16	0.01 (0.00–0.02)	0

Key: Disario et al, 2004[6]; Natsui et al, 2002[35]; Ochi et al, 1999[36]; Omar et al, 2017[37]; Oh and Kim, 2012[38]; Tanaka et al, 2004[44]; Arnold et al, 2001[45]; Lin et al, 2004[46]; Chu et al, 2016[50]; Fu et al, 2013[52]; Minami et al, 1995[54]; Guo et al, 2015.[53]

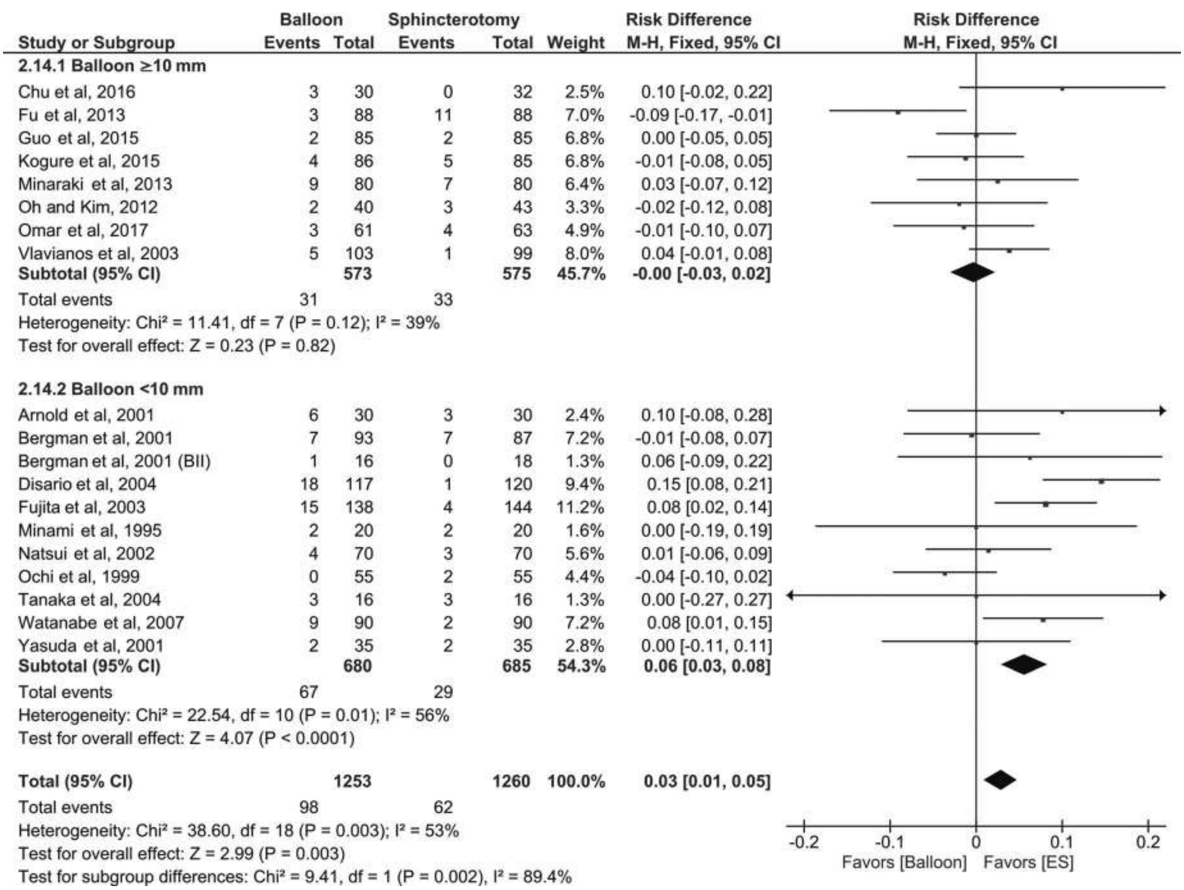

FIGURE 5. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of pancreatitis regarding balloon size.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.15.1 Asian countries						
Chu et al, 2016	3	30	0	32	2.2%	0.10 [-0.02, 0.22]
Fu et al, 2013	3	88	11	88	6.1%	-0.09 [-0.17, -0.01]
Fujita et al, 2003	15	138	4	144	9.8%	0.08 [0.02, 0.14]
Guo et al, 2015	2	85	2	85	5.9%	0.00 [-0.05, 0.05]
Kogure et al, 2015	4	86	5	85	6.0%	-0.01 [-0.08, 0.05]
Lin et al, 2004	0	51	0	53	3.6%	0.00 [-0.04, 0.04]
Minami et al, 1995	2	20	2	20	1.4%	0.00 [-0.19, 0.19]
Minaraki et al, 2013	9	80	7	80	5.6%	0.03 [-0.07, 0.12]
Natsui et al, 2002	4	70	3	70	4.9%	0.01 [-0.06, 0.09]
Ochi et al, 1999	0	55	2	55	3.8%	-0.04 [-0.10, 0.02]
Oh and Kim, 2012	2	40	3	43	2.9%	-0.02 [-0.12, 0.08]
Omar et al, 2017	3	61	4	63	4.3%	-0.01 [-0.10, 0.07]
Seo et al, 2013	5	62	5	60	4.3%	-0.00 [-0.10, 0.09]
Tanaka et al, 2004	3	16	3	16	1.1%	0.00 [-0.27, 0.27]
Watanabe et al, 2007	9	90	2	90	6.3%	0.08 [0.01, 0.15]
Yasuda et al, 2001	2	35	2	35	2.4%	0.00 [-0.11, 0.11]
Subtotal (95% CI)		**1007**		**1019**	**70.8%**	**0.01 [-0.01, 0.03]**
Total events	66		55			
Heterogeneity: Chi² = 22.14, df = 15 (P = 0.10); I² = 32%						
Test for overall effect: Z = 1.08 (P = 0.28)						
2.15.2 Non-Asian countries						
Arnold et al, 2001	6	30	3	30	2.1%	0.10 [-0.08, 0.28]
Bergman et al, 2001	7	93	7	87	6.3%	-0.01 [-0.08, 0.07]
Bergman et al, 2001 (BII)	1	16	0	18	1.2%	0.06 [-0.09, 0.22]
Disario et al, 2004	18	117	1	120	8.3%	0.15 [0.08, 0.21]
Omar et al, 2017	3	61	4	63	4.3%	-0.01 [-0.10, 0.07]
Vlavianos et al, 2003	5	103	1	99	7.1%	0.04 [-0.01, 0.08]
Subtotal (95% CI)		**420**		**417**	**29.2%**	**0.06 [0.02, 0.09]**
Total events	40		16			
Heterogeneity: Chi² = 12.87, df = 5 (P = 0.02); I² = 61%						
Test for overall effect: Z = 3.33 (P = 0.0009)						
Total (95% CI)		**1427**		**1436**	**100.0%**	**0.02 [0.01, 0.04]**
Total events	106		71			
Heterogeneity: Chi² = 41.09, df = 21 (P = 0.005); I² = 49%						
Test for overall effect: Z = 2.76 (P = 0.006)						
Test for subgroup differences: Chi² = 5.14, df = 1 (P = 0.02), I² = 80.5%						

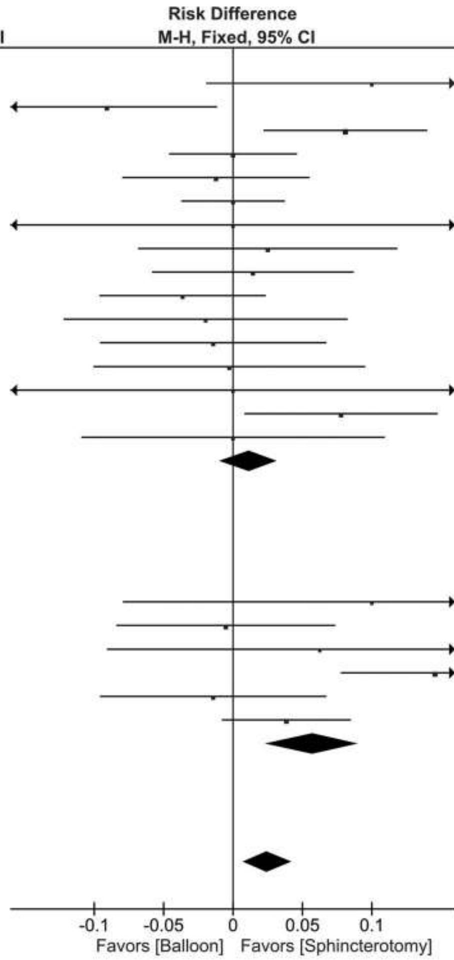

FIGURE 6. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of pancreatitis regarding ethnicity.

Comparison 2: Balloon Dilation Versus Balloon Dilation Plus ES

Four RCTs,[27,50,53,55] involving collective totals of 257 patients in the balloon dilation arms and 262 patients in the ES arms, were enrolled in this comparison and contributed to a pooled analysis of all outcomes except recurrence. Although stone removal in a single ERCP session was slightly higher in EPLBD (RD, −0.03; 95% CI, −0.09 to −0.02; $P = 0.25$), as was the need for ML (RD, 0.04; 95% CI, −0.01 to 0.09; $P = 0.12$), none of the differences were statistically significant. Adverse events were summarized as follows: pancreatitis (RD, −0.00; 95% CI, −0.04 to 0.03; $P = 0.88$); bleeding (RD, 0.00; 95% CI, −0.02 to 0.02; $P = 0.98$); cholangitis (RD, 0.00; 95% CI, −0.02 to 0.02; $P = 0.99$); and recurrence (RD, 0.00; 95% CI, −0.09 to 0.10; $P = 0.93$).

Heterogeneity was observed only for pancreatitis and for the need for ML, being less than substantial for both (I^2, 41% and I^2, 26%, respectively). All dilations were performed with balloons ≥10 mm, and all the studies provided information either of random sequence generation or allocation. Therefore, no subgroup analysis or sensitivity analysis regarding selection bias was carried out.

The results for pancreatitis were robust regarding ENBD (Table 3). All outcomes were rated as having moderate quality and were downgraded by 1 level for imprecision because the confidence intervals overlapped no effect, except for recurrence, given that the sample size was considered small and events were few (Table 4).

DISCUSSION

This is the first meta-analysis of RCT of balloon dilation in which PEP, one of the most worrisome adverse events in ERCP, was the main outcome measure. Historically, EPBD has been associated with higher pancreatitis rates especially after the publication of a multicenter RCT study conducted in the United States, in which more cases of severe pancreatitis (6 for EPBD vs 1 for ES) and 2 deaths were also reported.[6] In the initial overall analysis, our findings supported that association. However, it is believed that ES, standard balloon dilation, and EPLBD affect the sphincter of Oddi differently. A small balloon might not sufficiently dilate the papilla, and local edema occurs, obstructing the outflow of pancreatic secretion, predisposing to pancreatitis.[58] That scenario corresponds to our subgroup analysis, in which there were 67 cases of pancreatitis when dilation was performed with a balloon <10 mm versus 29 in the ES group. It is of note that all of the studies in this subgroup were older (published between 1995 and 2007) and that most of the dilations were performed with an 8 mm balloon even if the stones were >15 mm. The insufficient

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.18.1 Asian						
Chu et al, 2016	0	30	0	32	2.3%	0.00 [-0.06, 0.06]
Fu et al, 2013	0	88	0	88	6.6%	0.00 [-0.02, 0.02]
Fujita et al, 2003	0	138	0	144	10.6%	0.00 [-0.01, 0.01]
Kogure et al, 2015	0	86	0	85	6.4%	0.00 [-0.02, 0.02]
Lin et al, 2004	0	51	0	53	3.9%	0.00 [-0.04, 0.04]
Minami et al, 1995	0	20	0	20	1.5%	0.00 [-0.09, 0.09]
Minaraki et al, 2013	2	80	1	80	6.0%	0.01 [-0.03, 0.05]
Natsui et al, 2002	0	70	0	70	5.2%	0.00 [-0.03, 0.03]
Ochi et al, 1999	0	55	0	55	4.1%	0.00 [-0.03, 0.03]
Oh and Kim, 2012	0	40	0	43	3.1%	0.00 [-0.05, 0.05]
Seo et al, 2013	0	62	0	70	4.9%	0.00 [-0.03, 0.03]
Takezawa et al, 2004	0	46	0	45	3.4%	0.00 [-0.04, 0.04]
Tanaka et al, 2004	0	16	0	16	1.2%	0.00 [-0.11, 0.11]
Watanabe et al, 2007	1	90	0	90	6.7%	0.01 [-0.02, 0.04]
Yasuda et al, 2001	0	35	0	35	2.6%	0.00 [-0.05, 0.05]
Subtotal (95% CI)		**907**		**926**	**68.7%**	**0.00 [-0.01, 0.01]**
Total events	3		1			
Heterogeneity: Chi² = 0.84, df = 14 (P = 1.00); I² = 0%						
Test for overall effect: Z = 0.47 (P = 0.64)						
2.18.2 Non-asian						
Arnold et al, 2001	0	30	0	30	2.2%	0.00 [-0.06, 0.06]
Bergman et al, 2001	2	93	1	87	6.7%	0.01 [-0.03, 0.05]
Bergman et al, 2001 (BII)	0	16	0	18	1.3%	0.00 [-0.11, 0.11]
Disario et al, 2004	6	117	0	120	8.9%	0.05 [0.01, 0.09]
Omar et al, 2017	0	61	0	63	4.6%	0.00 [-0.03, 0.03]
Vlavianos et al, 2003	1	103	0	99	7.6%	0.01 [-0.02, 0.04]
Subtotal (95% CI)		**420**		**417**	**31.3%**	**0.02 [0.00, 0.04]**
Total events	9		1			
Heterogeneity: Chi² = 4.80, df = 5 (P = 0.44); I² = 0%						
Test for overall effect: Z = 2.09 (P = 0.04)						
Total (95% CI)		**1327**		**1343**	**100.0%**	**0.01 [-0.00, 0.02]**
Total events	12		2			
Heterogeneity: Chi² = 7.74, df = 20 (P = 0.99); I² = 0%						
Test for overall effect: Z = 1.75 (P = 0.08)						
Test for subgroup differences: Chi² = 2.72, df = 1 (P = 0.10), I² = 63.2%						

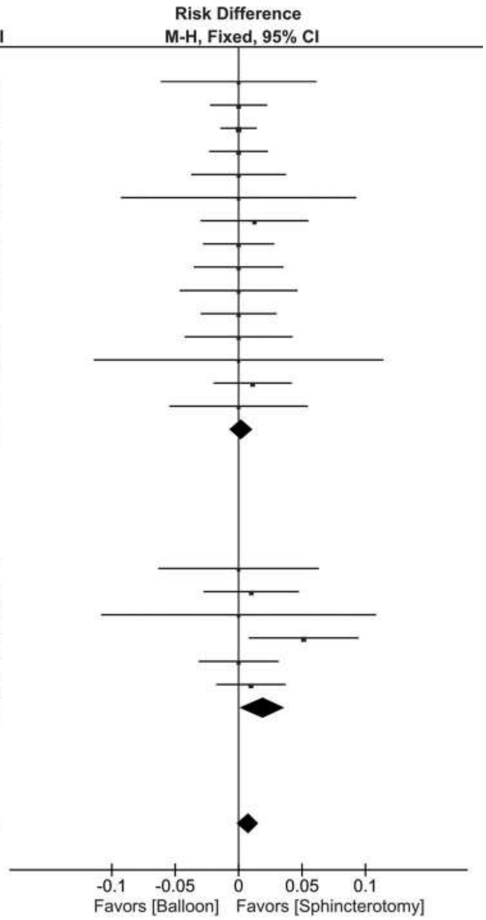

FIGURE 7. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of severe pancreatitis regarding ethnicity.

dilation and recent technological improvements could affect this result, lowering success rates and increasing the use of ML.

Another interesting analysis is based on the discussion of a recently published RCT. Park et al proposed that the divergent rates of PEP could derive from the higher prevalence of sphincter of Oddi dysfunction in non-Asian countries.[27] Our subgroup analysis showed that the incidence of pancreatitis was higher in non-Asian countries, indicating that this is an important issue to consider. In a similar vein, ENBD placement is a technique mostly used in Asian countries, sometimes replacing pancreatic stenting. Previous studies have shown improvement in PEP rates after balloon dilation and after ES + EPLBD,[59,60] although it remains controversial whether this can be beneficial after ES. In fact, in our analysis, when RCTs that reported this technique were omitted, the rates of pancreatitis and severe pancreatitis were higher in the balloon dilation arms.

The leave-one-out sensitivity analysis showed that the Disario et al[6] study had a major influence on the summary effect even among non-Asian countries and appeared to be a determinant of the between-study heterogeneity. This is likely because this trial was a well-conducted, high-quality RCT, which included a large sample size, and had an overall low risk of bias. Nevertheless, this study was limited by the fact that the sample size was younger than that in the other included studies. Seemingly, the incidence of sphincter of Oddi dysfunction is higher in younger individuals. Younger age has been identified as a risk factor for pancreatitis in many studies,[1,7] and the older cohort in those studies may under-represent the rates of pancreatitis in a younger population.[61]

It is currently recommended to dilate the papilla enough to accommodate the stone but not greater than the distal bile duct diameter, to avoid perforation. Larger balloons tend to dilate the papilla widely, avoiding the disadvantages of standard balloons and improving results.[17] When EPLBD was compared with ES, pancreatitis rates, ML use, and stone removal in the first ERCP were similar between the 2 techniques. Kogure et al[34] found that even when the stones were larger, balloon dilation decreased the need for ML and increased success rates, findings that run counter to those of the meta-analysis conducted by Kim et al,[12] as described in the current Japan Gastroenterological Endoscopy Society guidelines.[18]

International guidelines for EPLBD state that pancreatitis rates are similar between balloon dilation alone and ES + EPLBD.[17,18] In addition, the rate of stone removal in the first ERCP was lower when balloon dilation was performed, although the difference was not statistically significant. Our study of only RCTs (1A level of evidence) produced the same results, which was also similar to the previous network meta-analysis, when direct estimates were considered.[15] Since the publication of the meta-analysis conducted by Kim et al,[12] another 4 RCTs were published and included in our study.[27,37,39,50] In addition, recurrence and cholangitis were outcomes not reported in that previous article; in our analysis, the rates of both were similar.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.17.1 Balloon ≥10 mm						
Chu et al, 2016	0	30	0	32	2.4%	0.00 [-0.06, 0.06]
Fu et al, 2013	0	88	0	88	6.8%	0.00 [-0.02, 0.02]
Kogure et al, 2015	0	86	0	885	12.2%	0.00 [-0.02, 0.02]
Minaraki et al, 2013	2	80	1	80	6.2%	0.01 [-0.03, 0.05]
Oh and Kim, 2012	0	40	0	43	3.2%	0.00 [-0.05, 0.05]
Omar et al, 2017	0	61	0	63	4.8%	0.00 [-0.03, 0.03]
Vlavianos et al, 2003	1	103	0	99	7.8%	0.01 [-0.02, 0.04]
Subtotal (95% CI)		**488**		**1290**	**43.5%**	**0.00 [-0.01, 0.01]**
Total events	3		1			
Heterogeneity: Chi² = 0.75, df = 6 (P = 0.99); I² = 0%						
Test for overall effect: Z = 0.61 (P = 0.54)						
2.17.2 Balloon <10 mm						
Arnold et al, 2001	2	30	0	30	2.3%	0.07 [-0.04, 0.17]
Bergman et al, 2001	2	93	1	87	7.0%	0.01 [-0.03, 0.05]
Bergman et al, 2001 (BII)	0	16	0	18	1.3%	0.00 [-0.11, 0.11]
Disario et al, 2004	6	117	0	120	9.2%	0.05 [0.01, 0.09]
Fujita et al, 2003	0	138	0	144	10.9%	0.00 [-0.01, 0.01]
Minami et al, 1995	0	20	0	20	1.6%	0.00 [-0.09, 0.09]
Natsui et al, 2002	0	70	0	70	5.4%	0.00 [-0.03, 0.03]
Ochi et al, 1999	0	55	0	55	4.3%	0.00 [-0.03, 0.03]
Takezawa et al, 2004	0	46	0	45	3.5%	0.00 [-0.04, 0.04]
Tanaka et al, 2004	0	16	0	16	1.2%	0.00 [-0.11, 0.11]
Watanabe et al, 2007	1	90	0	90	7.0%	0.01 [-0.02, 0.04]
Yasuda et al, 2001	0	35	0	35	2.7%	0.00 [-0.05, 0.05]
Subtotal (95% CI)		**726**		**730**	**56.5%**	**0.01 [0.00, 0.03]**
Total events	11		1			
Heterogeneity: Chi² = 10.21, df = 11 (P = 0.51); I² = 0%						
Test for overall effect: Z = 2.11 (P = 0.03)						
Total (95% CI)		**1214**		**2020**	**100.0%**	**0.01 [0.00, 0.02]**
Total events	14		2			
Heterogeneity: Chi² = 10.31, df = 18 (P = 0.92); I² = 0%						
Test for overall effect: Z = 2.09 (P = 0.04)						
Test for subgroup differences: Chi² = 1.37, df = 1 (P = 0.24), I² = 27.0%						

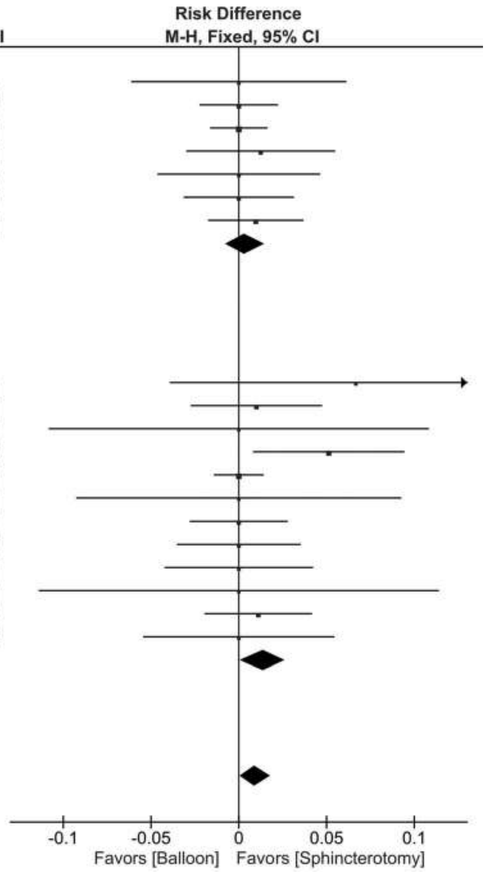

FIGURE 8. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of severe pancreatitis regarding balloon size.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Random, 95% CI
Arnold et al, 2001	23	30	30	30	4.4%	-0.23 [-0.39, -0.08]
Bergman et al, 2001	82	93	81	87	6.8%	-0.05 [-0.13, 0.04]
Bergman et al, 2001 (BII)	13	16	15	18	2.4%	-0.02 [-0.28, 0.24]
Chu et al, 2016	20	30	30	32	3.6%	-0.27 [-0.46, -0.08]
Disario et al, 2004	114	117	111	120	7.7%	0.05 [-0.01, 0.10]
Fu et al, 2013	99	103	82	103	6.7%	0.17 [0.08, 0.25]
Fujita et al, 2003	105	138	113	144	6.3%	-0.02 [-0.12, 0.07]
Guo et al, 2015	78	85	79	85	6.9%	-0.01 [-0.09, 0.07]
Kogure et al, 2015	78	86	67	85	6.0%	0.12 [0.01, 0.23]
Lin et al, 2004	41	51	47	53	4.9%	-0.08 [-0.22, 0.06]
Natsui et al, 2002	41	70	49	70	4.4%	-0.11 [-0.27, 0.04]
Ochi et al, 1999	40	55	51	55	5.0%	-0.20 [-0.34, -0.06]
Oh and Kim, 2012	33	40	35	43	4.2%	0.01 [-0.15, 0.18]
Omar et al, 2017	53	61	45	63	4.9%	0.15 [0.01, 0.29]
Seo et al, 2013	57	62	64	70	6.4%	0.01 [-0.09, 0.10]
Tanaka et al, 2004	11	16	9	16	1.6%	0.13 [-0.21, 0.46]
Vlavianos et al, 2003	65	103	63	99	5.1%	-0.01 [-0.14, 0.13]
Watanabe et al, 2007	65	90	52	90	5.0%	0.14 [0.01, 0.28]
Yasuda et al, 2001	35	35	35	35	7.7%	0.00 [-0.05, 0.05]
Total (95% CI)		**1281**		**1298**	**100.0%**	**-0.01 [-0.05, 0.04]**
Total events	1053		1058			
Heterogeneity: Tau² = 0.01; Chi² = 60.49, df = 18 (P < 0.00001); I² = 70%						
Test for overall effect: Z = 0.25 (P = 0.80)						

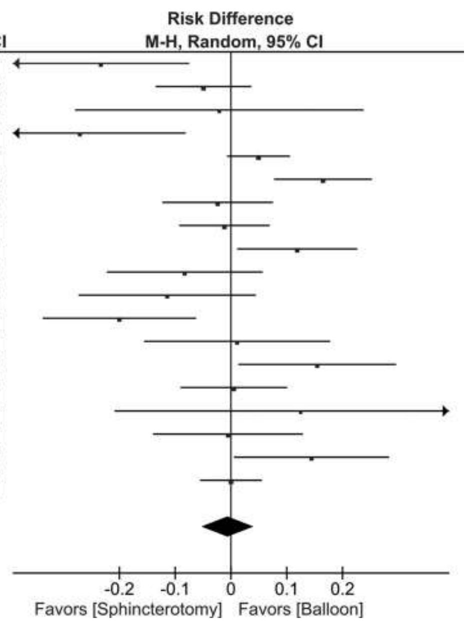

FIGURE 9. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of complete stone removal in the first ERCP.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Random, 95% CI
2.19.1 Balloon ≥10 mm						
Chu et al, 2016	20	30	26	32	3.4%	-0.15 [-0.36, 0.07]
Fu et al, 2013	99	103	82	103	7.6%	0.17 [0.08, 0.25]
Guo et al, 2015	78	85	79	85	7.9%	-0.01 [-0.09, 0.07]
Kogure et al, 2015	78	86	67	85	6.8%	0.12 [0.01, 0.23]
Oh and Kim, 2012	33	40	35	43	4.7%	0.01 [-0.15, 0.18]
Omar et al, 2017	53	61	45	63	5.5%	0.15 [0.01, 0.29]
Vlavianos et al, 2003	65	103	63	99	5.8%	-0.01 [-0.14, 0.13]
Subtotal (95% CI)		**508**		**510**	**41.7%**	**0.06 [-0.02, 0.13]**
Total events	426		397			
Heterogeneity: Tau² = 0.01; Chi² = 16.44, df = 6 (P = 0.01); I² = 63%						
Test for overall effect: Z = 1.46 (P = 0.14)						
2.19.2 Balloon <10 mm						
Arnold et al, 2001	23	30	30	30	4.9%	-0.23 [-0.39, -0.08]
Bergman et al, 2001	82	93	81	87	7.7%	-0.05 [-0.13, 0.04]
Bergman et al, 2001 (BII)	13	16	15	18	2.7%	-0.02 [-0.28, 0.24]
Disario et al, 2004	114	117	111	120	8.9%	0.05 [-0.01, 0.10]
Fujita et al, 2003	105	138	113	144	7.2%	-0.02 [-0.12, 0.07]
Natsui et al, 2002	41	70	49	70	4.9%	-0.11 [-0.27, 0.04]
Ochi et al, 1999	40	55	51	55	5.7%	-0.20 [-0.34, -0.06]
Tanaka et al, 2004	11	16	9	16	1.8%	0.13 [-0.21, 0.46]
Watanabe et al, 2007	65	90	52	90	5.6%	0.14 [0.01, 0.28]
Yasuda et al, 2001	35	35	35	35	8.9%	0.00 [-0.05, 0.05]
Subtotal (95% CI)		**660**		**665**	**58.3%**	**-0.03 [-0.10, 0.03]**
Total events	529		546			
Heterogeneity: Tau² = 0.01; Chi² = 28.44, df = 9 (P = 0.0008); I² = 68%						
Test for overall effect: Z = 1.06 (P = 0.29)						
Total (95% CI)		**1168**		**1175**	**100.0%**	**0.00 [-0.05, 0.05]**
Total events	955		943			
Heterogeneity: Tau² = 0.01; Chi² = 51.98, df = 16 (P < 0.0001); I² = 69%						
Test for overall effect: Z = 0.15 (P = 0.88)						
Test for subgroup differences: Chi² = 3.24, df = 1 (P = 0.07), I² = 69.1%						

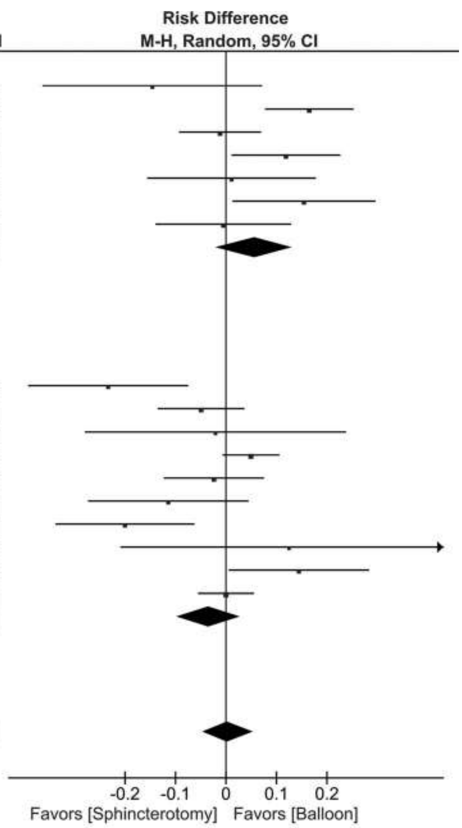

FIGURE 10. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of complete stone removal in the first ERCP regarding balloon size.

Our meta-analysis has some limitations. First, we carried out the analysis using RD to calculate effect sizes, because it focuses on the absolute effect. It is possible that different measurement choices tools might lead to other results, but these were not subsequently analyzed in this analysis. Second, the most recent consensus on EPLBD defines large balloons as those ≥12 mm and we

Study or Subgroup	Favors [Balloon] Events	Favors [Balloon] Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Random, 95% CI
Bergman et al, 2001	29	93	11	87	4.7%	0.19 [0.07, 0.30]
Bergman et al, 2001 (BII)	3	16	4	18	1.3%	-0.03 [-0.31, 0.24]
Chu et al, 2016	4	30	0	32	4.1%	0.13 [0.00, 0.26]
Disario et al, 2004	2	117	1	120	10.4%	0.01 [-0.02, 0.04]
Fu et al, 2013	3	103	4	103	9.1%	-0.01 [-0.06, 0.04]
Fujita et al, 2003	20	138	17	144	6.9%	0.03 [-0.05, 0.11]
Guo et al, 2015	12	85	8	85	5.8%	0.05 [-0.05, 0.14]
Kogure et al, 2015	26	86	41	85	3.7%	-0.18 [-0.32, -0.04]
Lin et al, 2004	1	51	2	53	8.0%	-0.02 [-0.08, 0.05]
Minaraki et al, 2013	2	80	2	80	9.1%	0.00 [-0.05, 0.05]
Natsui et al, 2002	29	70	27	70	3.1%	0.03 [-0.13, 0.19]
Ochi et al, 1999	7	55	2	55	5.6%	0.09 [-0.01, 0.19]
Oh and Kim, 2012	4	40	9	43	3.4%	-0.11 [-0.26, 0.04]
Omar et al, 2017	6	61	11	63	4.6%	-0.08 [-0.20, 0.04]
Seo et al, 2013	5	62	6	70	6.0%	-0.01 [-0.10, 0.09]
Tanaka et al, 2004	8	16	6	16	0.9%	0.13 [-0.22, 0.47]
Vlavianos et al, 2003	7	103	11	99	6.9%	-0.04 [-0.12, 0.04]
Watanabe et al, 2007	76	90	58	90	4.4%	0.20 [0.08, 0.32]
Yasuda et al, 2001	17	35	9	35	1.9%	0.23 [0.01, 0.45]
Total (95% CI)		**1331**		**1348**	**100.0%**	**0.02 [-0.01, 0.05]**
Total events	261		229			
Heterogeneity: Tau² = 0.00; Chi² = 45.59, df = 18 (P = 0.0003); I² = 61%						
Test for overall effect: Z = 1.19 (P = 0.24)						

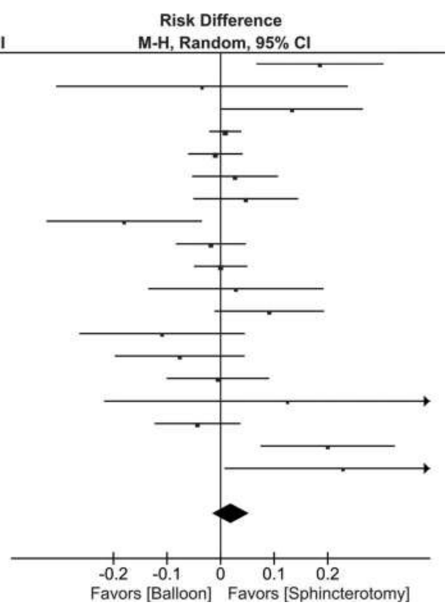

FIGURE 11. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of the need for ML.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Random, 95% CI
2.21.1 Balloon ≥10 mm						
Chu et al, 2016	4	30	0	32	5.0%	0.13 [0.00, 0.26]
Fu et al, 2013	3	103	4	103	10.1%	-0.01 [-0.06, 0.04]
Guo et al, 2015	12	85	8	85	6.8%	0.05 [-0.05, 0.14]
Kogure et al, 2015	26	86	41	85	4.4%	-0.18 [-0.32, -0.04]
Minaraki et al, 2013	2	80	2	80	10.2%	0.00 [-0.05, 0.05]
Oh and Kim, 2012	4	40	9	43	4.1%	-0.11 [-0.26, 0.04]
Omar et al, 2017	6	61	11	63	5.5%	-0.08 [-0.20, 0.04]
Vlavianos et al, 2003	7	103	11	99	8.0%	-0.04 [-0.12, 0.04]
Subtotal (95% CI)		**588**		**590**	**54.1%**	**-0.02 [-0.07, 0.03]**
Total events	64		86			
Heterogeneity: Tau² = 0.00; Chi² = 17.96, df = 7 (P = 0.01); I² = 61%						
Test for overall effect: Z = 0.80 (P = 0.43)						
2.21.2 Balloon <10 mm						
Bergman et al, 2001	29	93	11	87	5.7%	0.19 [0.07, 0.30]
Bergman et al, 2001 (BII)	3	16	4	18	1.7%	-0.03 [-0.31, 0.24]
Disario et al, 2004	2	117	1	120	11.4%	0.01 [-0.02, 0.04]
Fujita et al, 2003	20	138	17	144	8.0%	0.03 [-0.05, 0.11]
Natsui et al, 2002	29	70	27	70	3.8%	0.03 [-0.13, 0.19]
Ochi et al, 1999	7	55	2	55	6.6%	0.09 [-0.01, 0.19]
Tanaka et al, 2004	8	16	6	16	1.1%	0.13 [-0.22, 0.47]
Watanabe et al, 2007	76	90	58	90	5.3%	0.20 [0.08, 0.32]
Yasuda et al, 2001	17	35	9	35	2.4%	0.23 [0.01, 0.45]
Subtotal (95% CI)		**630**		**635**	**45.9%**	**0.09 [0.01, 0.18]**
Total events	191		135			
Heterogeneity: Tau² = 0.01; Chi² = 40.86, df = 8 (P < 0.00001); I² = 80%						
Test for overall effect: Z = 2.13 (P = 0.03)						
Total (95% CI)		**1218**		**1225**	**100.0%**	**0.03 [-0.01, 0.06]**
Total events	255		221			
Heterogeneity: Tau² = 0.00; Chi² = 44.53, df = 16 (P = 0.0002); I² = 64%						
Test for overall effect: Z = 1.35 (P = 0.18)						
Test for subgroup differences: Chi² = 5.00, df = 1 (P = 0.03), I² = 80.0%						

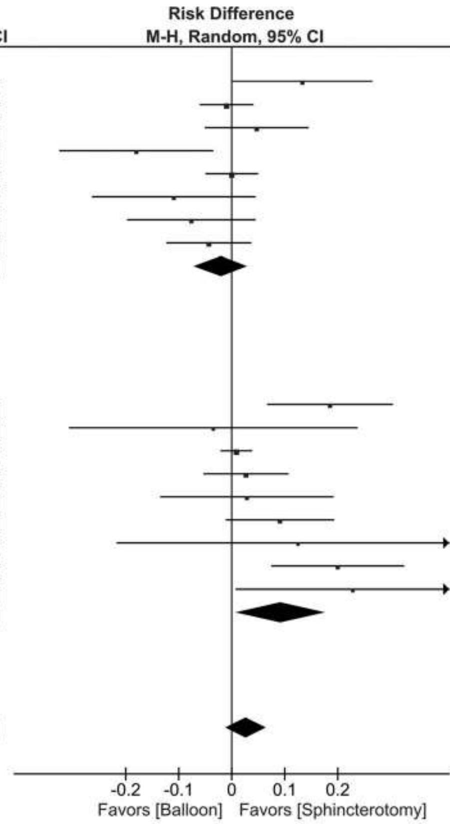

FIGURE 12. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of subgroup analysis of the use of ML regarding balloon size.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
Arnold et al, 2001	0	30	2	30	2.2%	-0.07 [-0.17, 0.04]
Bergman et al, 2001	0	93	2	87	6.7%	-0.02 [-0.06, 0.01]
Bergman et al, 2001 (BII)	0	16	3	18	1.3%	-0.17 [-0.36, 0.03]
Chu et al, 2016	0	30	2	32	2.3%	-0.06 [-0.16, 0.04]
Disario et al, 2004	0	117	0	120	8.9%	0.00 [-0.02, 0.02]
Fu et al, 2013	0	103	3	103	7.7%	-0.03 [-0.07, 0.01]
Fujita et al, 2003	0	138	2	144	10.6%	-0.01 [-0.04, 0.01]
Guo et al, 2015	1	85	1	85	6.4%	0.00 [-0.03, 0.03]
Kogure et al, 2015	0	86	1	85	6.4%	-0.01 [-0.04, 0.02]
Lin et al, 2004	1	51	0	53	3.9%	0.02 [-0.03, 0.07]
Minami et al, 1995	0	20	0	20	1.5%	0.00 [-0.09, 0.09]
Minaraki et al, 2013	1	80	1	80	6.0%	0.00 [-0.03, 0.03]
Natsui et al, 2002	0	70	2	70	5.2%	-0.03 [-0.08, 0.02]
Oh and Kim, 2012	0	40	0	43	3.1%	0.00 [-0.05, 0.05]
Omar et al, 2017	1	61	4	63	4.6%	-0.05 [-0.12, 0.02]
Seo et al, 2013	0	62	2	70	4.9%	-0.03 [-0.08, 0.02]
Tanaka et al, 2004	0	16	0	16	1.2%	0.00 [-0.11, 0.11]
Vlavianos et al, 2003	0	103	0	99	7.6%	0.00 [-0.02, 0.02]
Watanabe et al, 2007	0	90	1	90	6.7%	-0.01 [-0.04, 0.02]
Yasuda et al, 2001	0	35	1	35	2.6%	-0.03 [-0.10, 0.05]
Total (95% CI)		**1326**		**1343**	**100.0%**	**-0.02 [-0.03, -0.01]**
Total events	4		27			
Heterogeneity: Chi² = 17.79, df = 19 (P = 0.54); I² = 0%						
Test for overall effect: Z = 3.35 (P = 0.0008)						

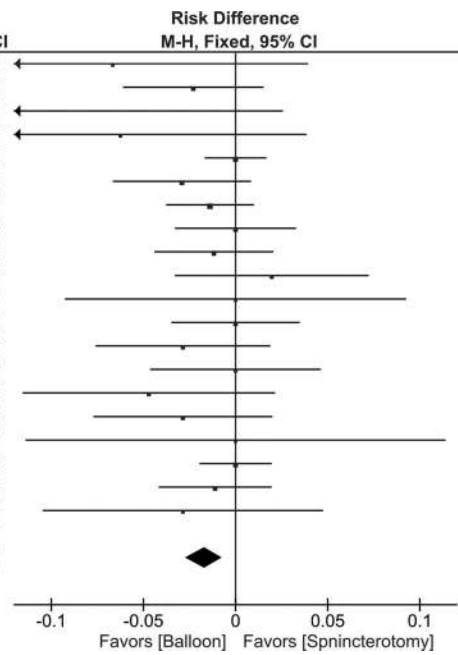

FIGURE 13. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of bleeding.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
Arnold et al, 2001	3	30	0	30	3.1%	0.10 [-0.02, 0.22]
Chu et al, 2016	0	30	0	32	3.2%	0.00 [-0.06, 0.06]
Disario et al, 2004	1	117	1	120	12.2%	0.00 [-0.02, 0.02]
Fu et al, 2013	0	103	0	103	10.6%	0.00 [-0.02, 0.02]
Guo et al, 2015	1	85	1	85	8.8%	0.00 [-0.03, 0.03]
Kogure et al, 2015	3	86	2	85	8.8%	0.01 [-0.04, 0.06]
Natsui et al, 2002	2	70	3	70	7.2%	-0.01 [-0.08, 0.05]
Ochi et al, 1999	0	55	0	55	5.7%	0.00 [-0.03, 0.03]
Oh and Kim, 2012	2	40	1	43	4.3%	0.03 [-0.05, 0.11]
Omar et al, 2017	2	61	1	63	6.4%	0.02 [-0.04, 0.07]
Takezawa et al, 2004	0	46	0	45	4.7%	0.00 [-0.04, 0.04]
Tanaka et al, 2004	0	16	2	16	1.7%	-0.13 [-0.31, 0.06]
Vlavianos et al, 2003	2	103	1	99	10.4%	0.01 [-0.02, 0.04]
Watanabe et al, 2007	3	90	0	90	9.3%	0.03 [-0.01, 0.08]
Yasuda et al, 2001	0	35	0	35	3.6%	0.00 [-0.05, 0.05]
Total (95% CI)		**967**		**971**	**100.0%**	**0.01 [-0.01, 0.02]**
Total events	19		12			

Heterogeneity: Chi² = 8.08, df = 14 (P = 0.89); I² = 0%
Test for overall effect: Z = 1.16 (P = 0.25)

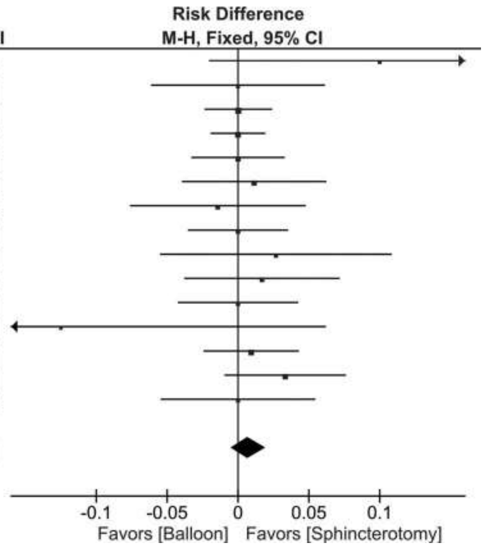

FIGURE 14. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of acute cholangitis.

considered balloons 10 mm or greater eligible, to embrace more studies. Another potential limitation was the number of studies conducted in non-Asian countries. The only RCTs comparing EPLBD with the other 2 techniques were conducted in Asian countries. Previous discussions have also questioned whether both groups are even comparable, given their peculiarities. Another Additionally, it is known that there are other diverse procedure-related PEP risk factors, such as difficult cannulation, guidewire cannulation, and/or pancreatic duct contrast injection, and most RCTs did not provide sufficient information to investigate these risk factors and their effects on PEP further.[62] In addition, only a few of the studies reported data on the use of nonsteroidal anti-inflammatory drugs or pancreatic stenting for PEP prevention.[63–65] Notably, in 2 studies,[6,49] pancreatography was performed in more than 50% of the patients. Finally, although recurrence was similar in both comparisons, there is a lack of long-term RCTs on the theme and each study had a different follow-up duration.

The important issues raised above regarding PEP indicate that analysis of this adverse event should consider its multiple risk factors, not all of which could be controlled for equally in all of the studies. In addition, most of the EPLBD studies were conducted in Asian countries and selected older participants, reducing the external validity of the results.

CONCLUSIONS

The incidence of PEP was similar among EPLBD, ES, and ES + EPLBD. In addition, EPLBD appears to be a safe alternative for the removal of large bile duct stones, given that there was also no difference among the techniques in terms of the need for ML, the stone removal rate, and the recurrence rate. Nevertheless, the results should be interpreted with caution, because pancreatitis has multiple risk factors which should be weighted as confounders and RCTs can have limited generalizability.

Study or Subgroup	Balloon Events	Balloon Total	Sphincterotomy Events	Sphincterotomy Total	Weight	Risk Difference M-H, Fixed, 95% CI
Bergman et al, 1997	8	101	7	101	13.1%	0.01 [-0.06, 0.08]
Chu et al, 2016	0	30	2	32	4.0%	-0.06 [-0.16, 0.04]
Disario et al, 2004	9	117	3	120	15.4%	0.05 [-0.00, 0.11]
Lin et al, 2004	3	51	4	53	6.7%	-0.02 [-0.11, 0.08]
Natsui et al, 2002	3	70	3	70	9.1%	0.00 [-0.07, 0.07]
Ochi et al, 1999	2	55	3	55	7.1%	-0.02 [-0.10, 0.06]
Omar et al, 2017	2	61	2	63	8.0%	0.00 [-0.06, 0.06]
Seo et al, 2013	1	62	4	70	8.5%	-0.04 [-0.10, 0.02]
Takezawa et al, 2004	3	46	1	45	5.9%	0.04 [-0.04, 0.13]
Tanaka et al, 2004	5	35	5	35	4.5%	0.00 [-0.16, 0.16]
Vlavianos et al, 2003	2	103	3	99	13.1%	-0.01 [-0.05, 0.03]
Yasuda et al, 2001	2	35	3	35	4.5%	-0.03 [-0.15, 0.09]
Total (95% CI)		**766**		**778**	**100.0%**	**0.00 [-0.02, 0.02]**
Total events	40		40			

Heterogeneity: Chi² = 8.36, df = 11 (P = 0.68); I² = 0%
Test for overall effect: Z = 0.07 (P = 0.95)

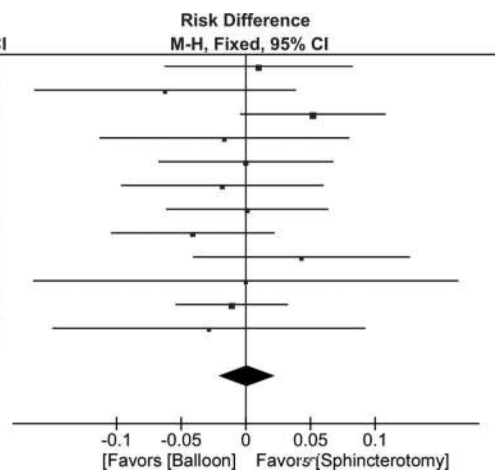

FIGURE 15. Meta-analysis of trials comparing balloon dilation and sphincterotomy. Forest plot of recurrence.

TABLE 4. Balloon Dilation Versus Balloon Dilation Plus Sphincterotomy

No. Studies	No. Patients, n (%)		Effect			
	Balloon	ES + Balloon	RD (95% CI)	*P*	I^2	GRADE
Pancreatitis						
4	9/257 (3.5)	10/262 (3.8)	−0.00 (−0.04 to 0.03)	0.88	26%	⊕⊕⊕○ Moderate*
Stone removal in 1 session						
4	216/257 (84.0)	229/262 (87.4)	−0.03 (−0.09 to 0.02)	0.25	0%	⊕⊕⊕○ Moderate*
Lithotripsy						
4	31/257 (12.1)	21/262 (8.0)	0.04 (−0.01 to 0.09)	0.12	41%	⊕⊕⊕⊕ High
Cholangitis						
3	2/227 (0.9)	2/229 (0.9)	0.00 (−0.02 to 0.02)	0.99	0%	⊕⊕⊕○ Moderate*
Bleeding						
4	2/257 (0.8)	2/262 (0.8)	0.00 (−0.02 to 0.02)	0.98	0%	⊕⊕⊕○ Moderate*
Recurrence						
2	7/72 (9.7)	7/77 (9.1)	0.00 (−0.09 to 0.10)	0.93	0%	⊕⊕○○ Low*†

*Downgrade by 1 level for imprecision because the confidence interval overlaps no effect.
†Downgrade by 1 level for imprecision due to a small sample size and few events.

REFERENCES

1. Leerhøy B, Elmunzer BJ. How to avoid post-endoscopic retrograde cholangiopancreatography pancreatitis. *Gastrointest Endosc Clin N Am*. 2018;28:439–454.
2. Furuya CK, Sakai P, Marinho FRT, et al. Papillary fistulotomy vs conventional cannulation for endoscopic biliary access: a prospective randomized trial. *World J Gastroenterol*. 2018;24: 1803–1811.
3. de Clemente Junior CC, Bernardo WM, Franzini TP, et al. Comparison between endoscopic sphincterotomy vs endoscopic sphincterotomy associated with balloon dilation for removal of bile duct stones: a systematic review and meta-analysis based on randomized controlled trials. *World J Gastrointest Endosc*. 2018;10:130–144.
4. Franzini T, Moura RN, Bonifácio P, et al. Complex biliary stones management: cholangioscopy versus papillary large balloon dilation - a randomized controlled trial. *Endosc Int Open*. 2018;6: E131–E138.
5. Staritz M, Ewe K, Meyer zum Büschenfelde K. Endoscopic papillary dilation (EPD) for the treatment of common bile duct stones and papillary stenosis. *Endoscopy*. 1983;15(Suppl 1):197–198.
6. Disario JA, Freeman ML, Bjorkman DJ, et al. Endoscopic balloon dilation compared with sphincterotomy for extraction of bile duct stones. *Gastroenterology*. 2004;127:1291–1299.
7. Freeman ML, Disario JA, Nelson DB, et al. Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. *Gastrointest Endosc*. 2001;54:425–434.
8. Weinberg BM, Shindy W, Lo S. Endoscopic balloon sphincter dilation (sphincteroplasty) versus sphincterotomy for common bile duct stones. *Cochrane Database Syst Rev*. 2006;CD004890.
9. Baron TH, Harewood GC. Endoscopic balloon dilation of the biliary sphincter compared to endoscopic biliary Sphincterotomy for removal of common bile duct stones during ERCP: a meta-analysis of randomized, controlled trials. *Am J Gastroenterol*. 2004;99:1455–1460.
10. Ersoz G, Tekesin O, Ozutemiz AO, et al. Biliary sphincterotomy plus dilation with a large balloon for bile duct stones that are difficult to extract. *Gastrointest Endosc*. 2003;57:156–159.
11. Meine GC, Baron TH. Endoscopic papillary large-balloon dilation combined with endoscopic biliary sphincterotomy for the removal of bile duct stones (with video). *Gastrointest Endosc*. 2011;74:1119–1126; quiz 1115.e1–e5.
12. Kim JH, Yang MJ, Hwang JC, et al. Endoscopic papillary large balloon dilation for the removal of bile duct stones. *World J Gastroenterol*. 2013;19: 8580–8594.
13. Jeong S, Ki SH, Lee DH, et al. Endoscopic large-balloon sphincteroplasty without preceding sphincterotomy for the removal of large bile duct stones : a preliminary study. *Gastrointest Endosc*. 2009;70:915–922.
14. Xu L, Kyaw MH, Tse YK, et al. Endoscopic Sphincterotomy with large balloon dilation versus endoscopic Sphincterotomy for bile duct stones: a systematic review and meta-analysis. *Biomed Res Int*. 2015; 2015:673103.
15. Park CH, Jung JH, Nam E, et al. Comparative efficacy of various endoscopic techniques for the treatment of common bile duct stones: a network meta-analysis. *Gastrointest Endosc*. 2018;87:43–57.e10.
16. Fujisawa T, Kagawa K, Hisatomi K, et al. Is endoscopic papillary balloon dilatation really a risk factor for post-ERCP pancreatitis? *World J Gastroenterol*. 2016;22:5909–5916.
17. Kim TH, Kim JH, Seo DW, et al. International consensus guidelines for endoscopic papillary large-balloon dilation. *Gastrointest Endosc*. 2016;83: 37–47.
18. Itoi T, Ryozawa S, Katanuma A, et al. Japan Gastroenterological Endoscopy Society guidelines for endoscopic papillary large balloon dilation. *Dig Endosc*. 2018;30:293–309.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264–269, W64.

20. Cotton PB, Lehman G, Vennes J, et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. *Gastrointest Endosc*. 1991;37:383–393.
21. Schünemann H, Brożek J, Guyatt G, et al. GRADE handbook for grading quality of evidence and strength of recommendations. *The GRADE Working Group*. 2013. Updated October 2013. Available at: https://gdt.gradepro.org/app/handbook/handbook.html. Accessed May 10, 2019.
22. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22:719–748.
23. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
24. Higgins JGS. Cochrane handbook for systematic reviews of interventions version 5.1.0, chapter 9. [online]. *Cochrane Collab*. 2011. Updated March 2011. Available at: https://training.cochrane.org/handbook. Accessed May 10, 2019.
25. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177–188.
26. Patsopoulos NA, Evangelou E, Ioannidis JP, et al. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. *Int J Epidemiol*. 2008;37:1148–1157.
27. Park JS, Jeong S, Lee DK, et al. Comparison of endoscopic papillary large balloon dilation with or without endoscopic sphincterotomy for the treatment of large bile duct stones. *Endoscopy*. 2019;51:125–132.
28. Iwata F, Miyaki T, Yamada J, et al. Effect of endoscopic papillary balloon dilation for the management of common bile duct stones in the elderly. *Gastroenterology*. 1988;144(suppl 1):A524.abstract G2138.
29. Chen YK, Deguzman L, Godil A, et al. Prospective randomized trial of endoscopic balloon dilation versus sphincterotomy for extraction of bile duct stones. *Gastrointest Endosc*. 1998;47:AB111. abstract 347.
30. Cho YD, Hong SJ, Moon JH, et al. Endoscopic papillary balloon dilatation versus endoscopic sphincterotomy for removal of common bile duct stones. *Gastrointest Endosc*. 1998;47:AB111.abstract 349.
31. Sarhan M, Enaba M, El-Bedewy M, et al. Is combination of biliary sphincterotomy and balloon dilation a better option than either alone in endoscopic removal of bile duct stones? A comparative study. *J Gastroenterol Hepatol (Australia)*. 2014;29(suppl 3):154. abstract P-299.
32. Na H, Chun Y, Lee T, et al. Sphinter of Oddi function after endoscopic papillary balloon dilation alone vs. endoscopic papillary balloon dilation with endoscopic sphincterotomy. *J Gastroenterol Hepatol*. 2016; 31(suppl 3):328.abstract 928.
33. Li CM, Liu ZH, Jiang GL, et al. Clinical evaluation of endoscopic sphincterotomy and papillary balloon dilation for removal of common bile duct stones. *World Chinese J Dig*. 2006;14:230–233.
34. Kogure H, Kawahata S, Mukai T, et al. A multicenter randomized trial of endoscopic papillary large balloon dilation alone versus endoscopic sphincterotomy for removal of bile duct stones: MARVELOUS trial. *United Eur Gastroenterol J*. 2015;3(suppl 1):A69.abstract OP211.
35. Natsui M, Narisawa R, Motoyama H, et al. What is an appropriate indication for endoscopic papillary balloon dilation? *Eur J Gastroenterol Hepatol*. 2002;14:635–640.
36. Ochi Y, Mukawa K, Kiyosawa K, et al. Comparing the treatment outcomes of endoscopic papillary dilation and endoscopic sphincterotomy for removal of bile duct stones. *J Gastroenterol Hepatol*. 1999;14:90–96.
37. Omar MA, Abdelshafy M, Ahmed MY, et al. Endoscopic papillary large balloon dilation versus endoscopic sphincterotomy for retrieval of large choledocholithiasis: a prospective randomized trial. *J Laparoendosc Adv Surg Tech A*. 2017;27:704–709.
38. Oh MJ, Kim TN. Prospective comparative study of endoscopic papillary large balloon dilation and endoscopic sphincterotomy for removal of large bile duct stones in patients above 45 years of age. *Scand J Gastroenterol*. 2012;47:1071–1077.
39. Minakari M, Samani RR, Shavakhi A, et al. Endoscopic papillary balloon dilatation in comparison with endoscopic sphincterotomy for the treatment of large common bile duct stone. *Adv Biomed Res*. 2013;2:46.
40. Yasuda I, Tomita E, Enya M, et al. Can endoscopic papillary balloon dilation really preserve sphincter of Oddi function? *Gut*. 2001;49: 686–691.
41. Vlavianos P, Chopra K, Mandalia S, et al. Endoscopic balloon dilatation versus endoscopic sphincterotomy for the removal of bile duct stones: a prospective randomised trial. *Gut*. 2003;52:1165–1169.
42. Watanabe H, Yoneda M, Tominaga K, et al. Comparison between endoscopic papillary balloon dilatation and endoscopic sphincterotomy for the treatment of common bile duct stones. *J Gastroenterol*. 2007;42:56–62.
43. Takezawa M, Kida Y, Kida M, et al. Influence of endoscopic papillary balloon dilation and endoscopic sphincterotomy on sphincter of Oddi function: a randomized controlled trial. *Endoscopy*. 2004;36: 631–637.
44. Tanaka S, Sawayama T, Yoshioka T. Endoscopic papillary balloon dilation and endoscopic sphincterotomy for bile duct stones: long-term outcomes in a prospective randomized controlled trial. *Gastrointest Endosc*. 2004;59: 614–618.
45. Arnold JC, Benz C, Martin WR, et al. Endoscopic papillary balloon dilation vs. sphincterotomy for removal of common bile duct stones: a prospective randomized pilot study. *Endoscopy*. 2001;33: 563–567.
46. Lin CK, Lai KH, Chan HH, et al. Endoscopic balloon dilatation is a safe method in the management of common bile duct stones. *Dig Liver Dis*. 2004;36:68–72.
47. Seo YR, Moon JH, Choi HJ, et al. Comparison of endoscopic papillary balloon dilation and sphincterotomy in young patients with CBD stones and gallstones. *Dig Dis Sci*. 2014;59:1042–1047.
48. Bergman JJ, Van Berkel AM, Bruno MJ, et al. Is endoscopic balloon dilation for removal of bile duct stones associated with an increased risk for pancreatitis or a higher rate of hyperamylasemia? *Endoscopy*. 2001;33: 416–420.
49. Bergman JJ, Van Berkel AM, Bruno MJ, et al. A randomized trial of endoscopic balloon dilation and endoscopic sphincterotomy for removal of bile duct stones in patients with a prior Billroth II gastrectomy. *Gastrointest Endosc*. 2001;53:19–26.
50. Chu X, Zhang H, Qu R, et al. Small endoscopic sphincterotomy combined with endoscopic papillary large-balloon dilation in the treatment of patients with large bile duct stones. *Eur Surg*. 2017;49:9–16.
51. Fujita N, Maguchi H, Komatsu Y, et al. Endoscopic sphincterotomy and endoscopic papillary balloon dilatation for bile duct stones: a prospective randomized controlled multicenter trial. *Gastrointest Endosc*. 2003;57: 151–155.
52. Fu BQ, Xu YP, Tao LS, et al. Endoscopic papillary balloon intermittent dilatation and endoscopic sphincterotomy for bile duct stones. *World J Gastroenterol*. 2013;19:2425–2432.
53. Guo Y, Lei S, Gong W, et al. A preliminary comparison of endoscopic Sphincterotomy, endoscopic papillary large balloon dilation, and combination of the two in endoscopic choledocholithiasis treatment. *Med Sci Monit*. 2015;21:2607–2612.
54. Minami A, Nakatsu T, Uchida N, et al. Papillary dilation vs sphincterotomy in endoscopic removal of bile duct stones. A randomized trial with manometric function. *Dig Dis Sci*. 1995;40:2550–2554.
55. Cheon YK, Lee TY, Kim SN, et al. Impact of endoscopic papillary large-balloon dilation on sphincter of Oddi function: a prospective randomized study. *Gastrointest Endosc*. 2017;85:782–790.e1.
56. Guo Y, Li C, Lei S, et al. Effects comparison between endoscopic papillary large balloon dilatation and endoscopic sphincterotomy for common bile duct stone removal. *Gastroenterol Res Pract*. 2015;2015:839346.

57. Bergman JJ, Rauws EA, Fockens P, et al. Randomised trial of endoscopic balloon dilation versus endoscopic sphincterotomy for removal of bile duct stones. *Lancet*. 1997;349:1124–1129.
58. Ding J, Li F, Zhu HY, et al. Endoscopic treatment of difficult extrahepatic bile duct stones, EPBD or EST: an anatomic view. *World J Gastrointest Endosc*. 2015;7:274–277.
59. Sato D, Shibahara T, Miyazaki K, et al. Efficacy of endoscopic nasobiliary drainage for the prevention of pancreatitis after papillary balloon dilatation. *Pancreas*. 2005;31:93–97.
60. Xu X, Dai J, Qian J, et al. Nasobiliary drainage after endoscopic papillary balloon dilatation may prevent postoperative pancreatitis. *World J Gastroenterol*. 2015;21:2443–2449.
61. Disario JA. Endoscopic balloon dilation for extraction of bile duct stones: the devil is in the details. *Gastrointest Endosc*. 2003;57:282–285.
62. Serrano JPR, de Moura DTH, Bernardo WM, et al. Nonsteroidal anti-inflammatory drugs versus placebo for post-endoscopic retrograde cholangiopancreatography pancreatitis: a systematic review and meta-analysis. *Endosc Int Open*. 2019;7:E477–E486.
63. ASGE Standards of Practice Committee, Chandrasekhara V, Khashab MA, et al. Adverse events associated with ERCP. *Gastrointest Endosc*. 2017;85:32–47.
64. Testoni P, Mariani A, Aabakken L, et al. Papillary cannulation and sphincterotomy techniques at ERCP: European Society of Gastrointestinal Endoscopy (ESGE) clinical guideline. *Endoscopy*. 2016;48:657–683.
65. Cennamo V, Fuccio L, Zagari RM, et al. Can a wire-guided Cannulation technique increase bile duct Cannulation rate and prevent post-ERCP pancreatitis ?: a meta-analysis of randomized controlled trials. *Am J Gastroenterol*. 2009;104:2343–2350.