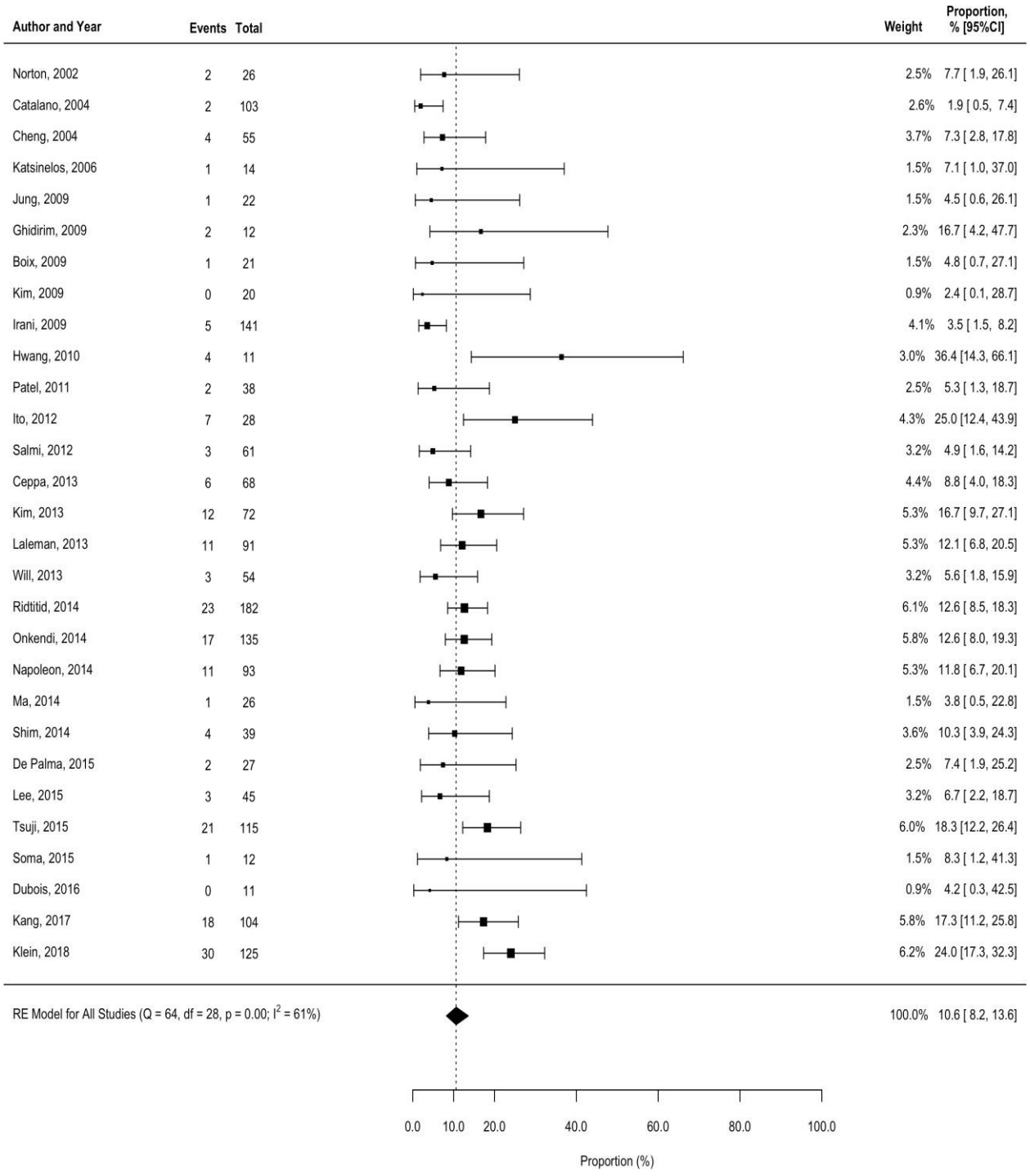
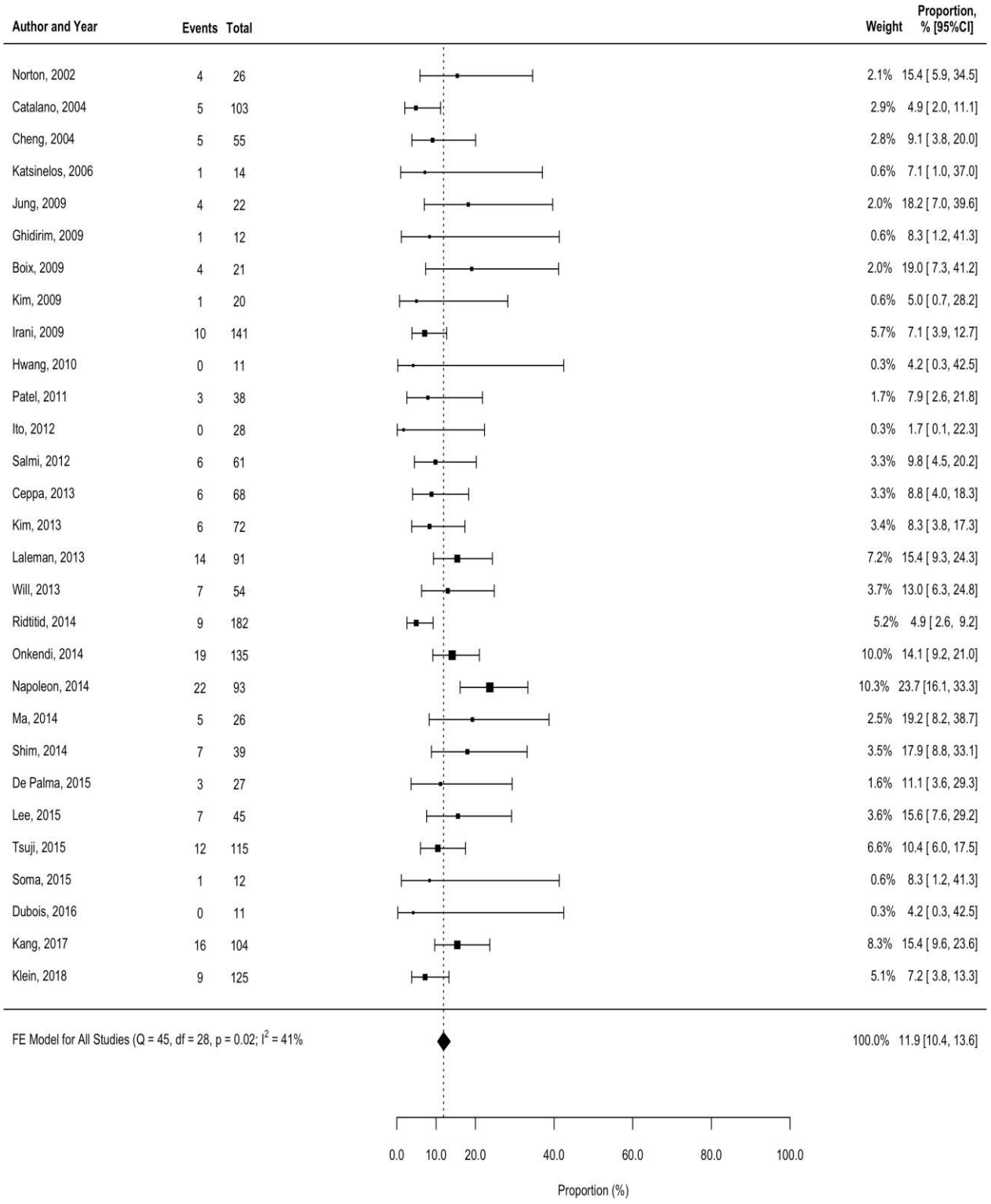
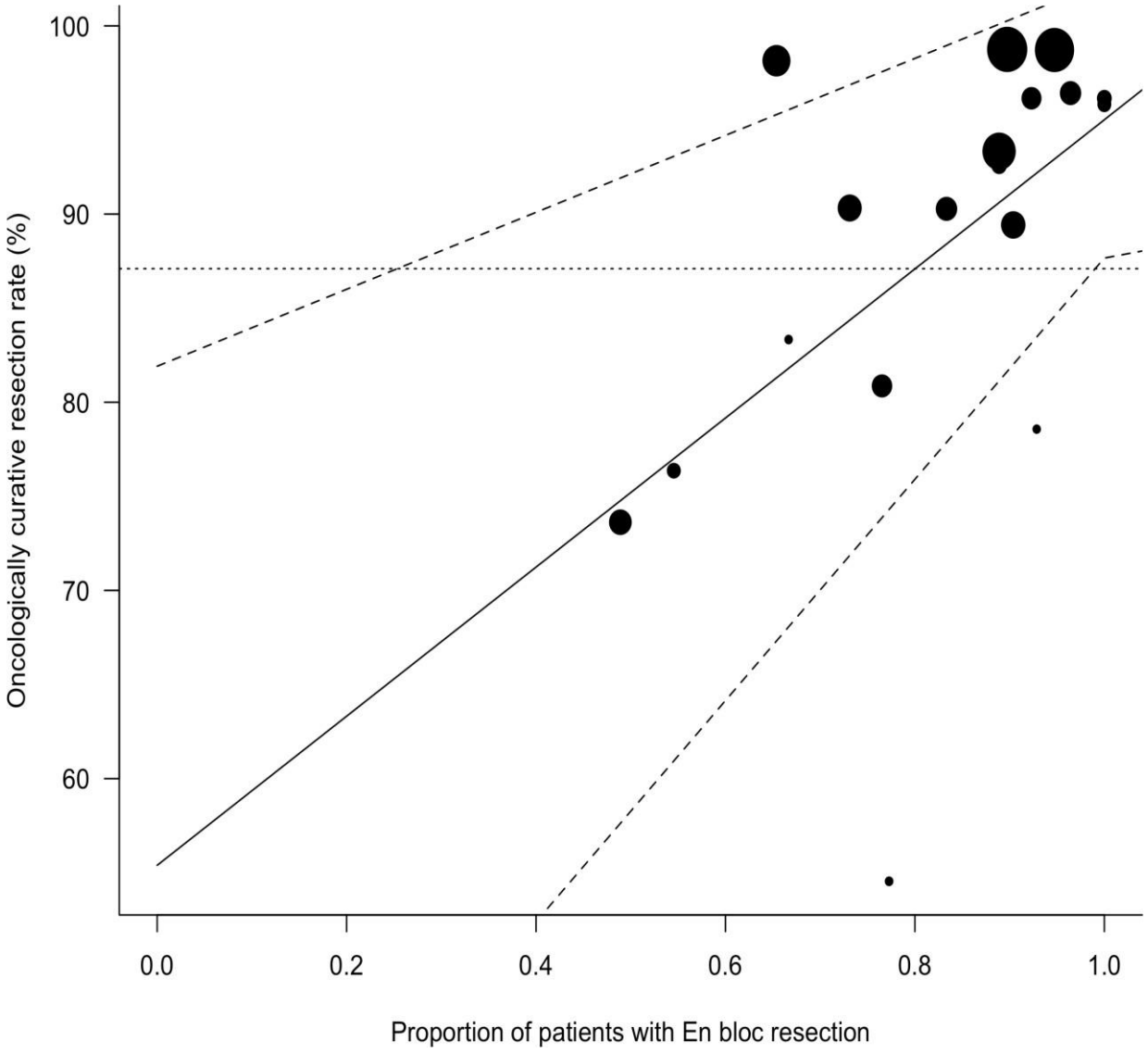
Supplementary figure 1: Forest plot reporting the rates of bleedings. CI: confidence interval. RE: random effect.



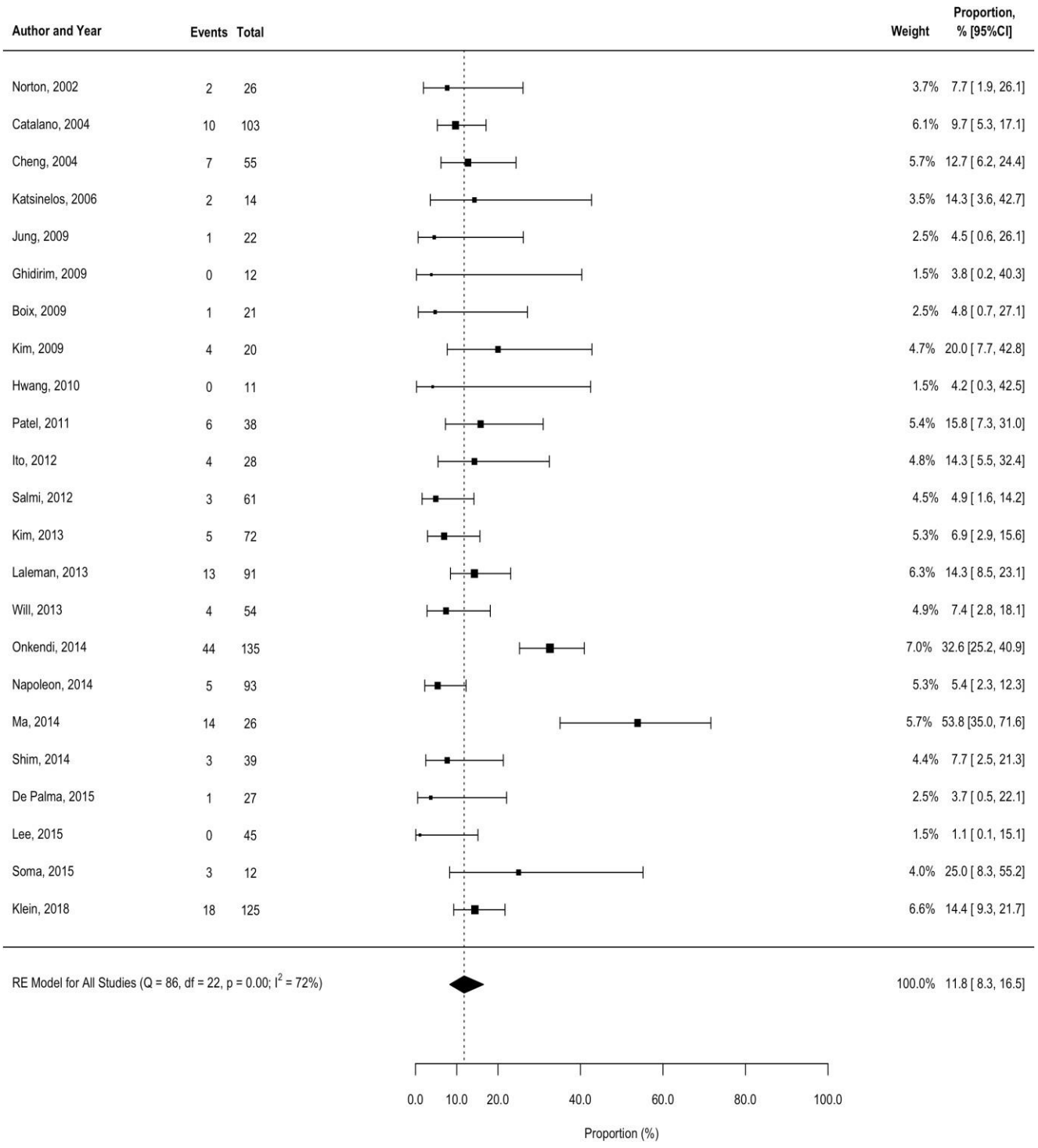
Supplementary figure 2: Forest plot reporting the rates of pancreatitis. CI: confidence interval. RE: random effect.



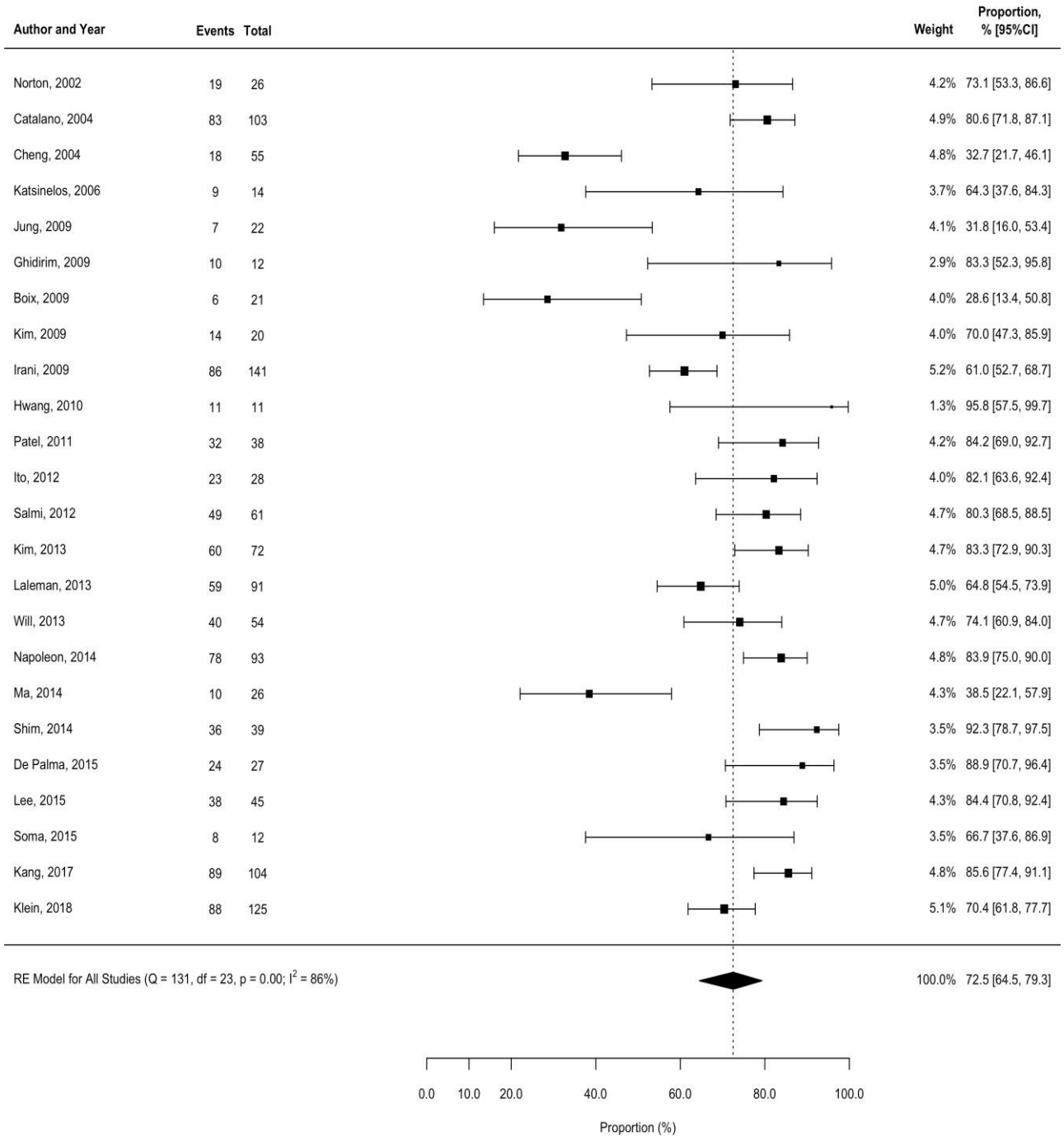
Supplementary figure 3: Bubble plot reporting the association between the curative resection occurrence and the proportion of patients with en-bloc resection.



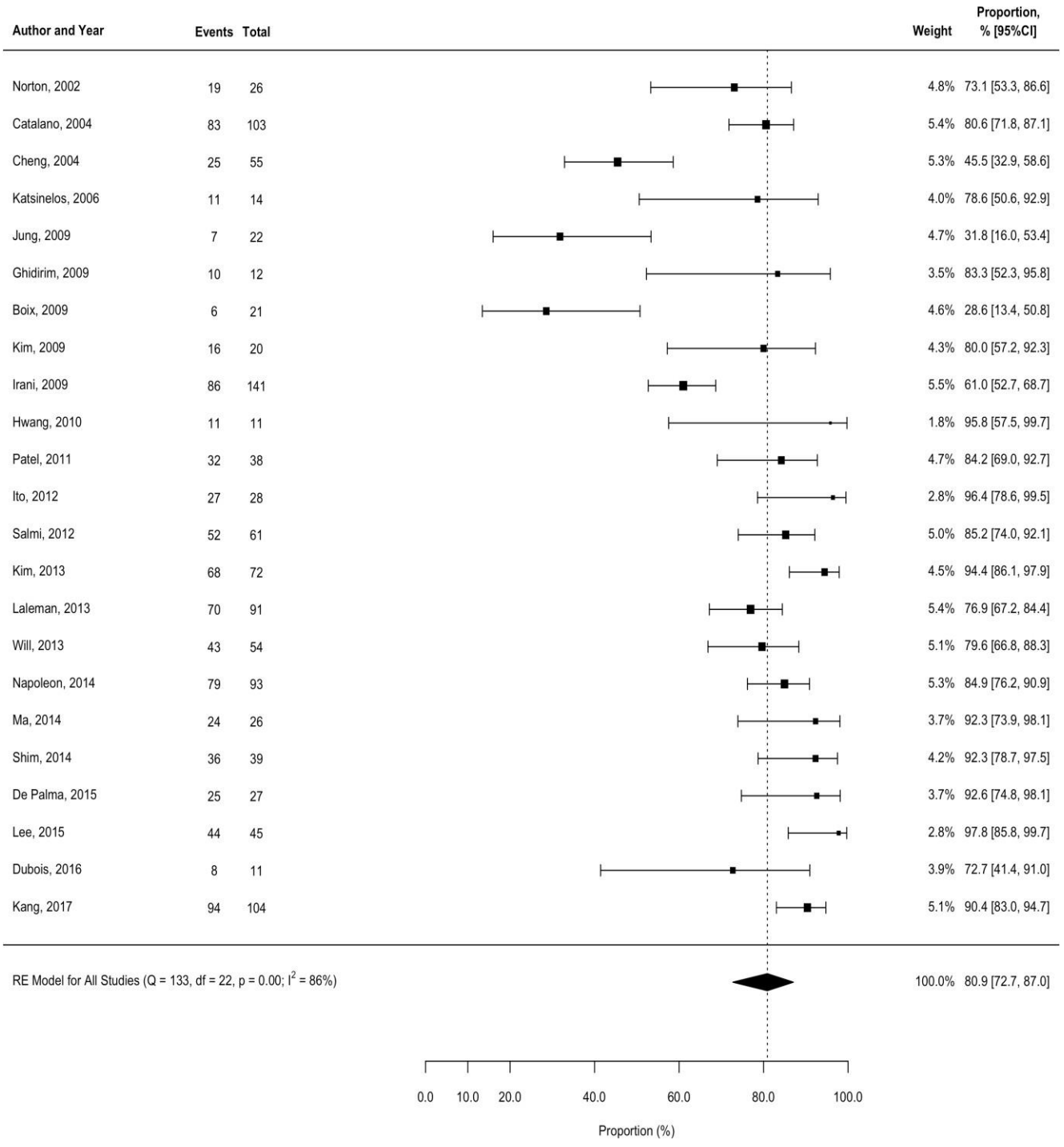
Supplementary figure 4: Forest plot reporting the rates of recurrence. CI: confidence interval. RE: random effect.



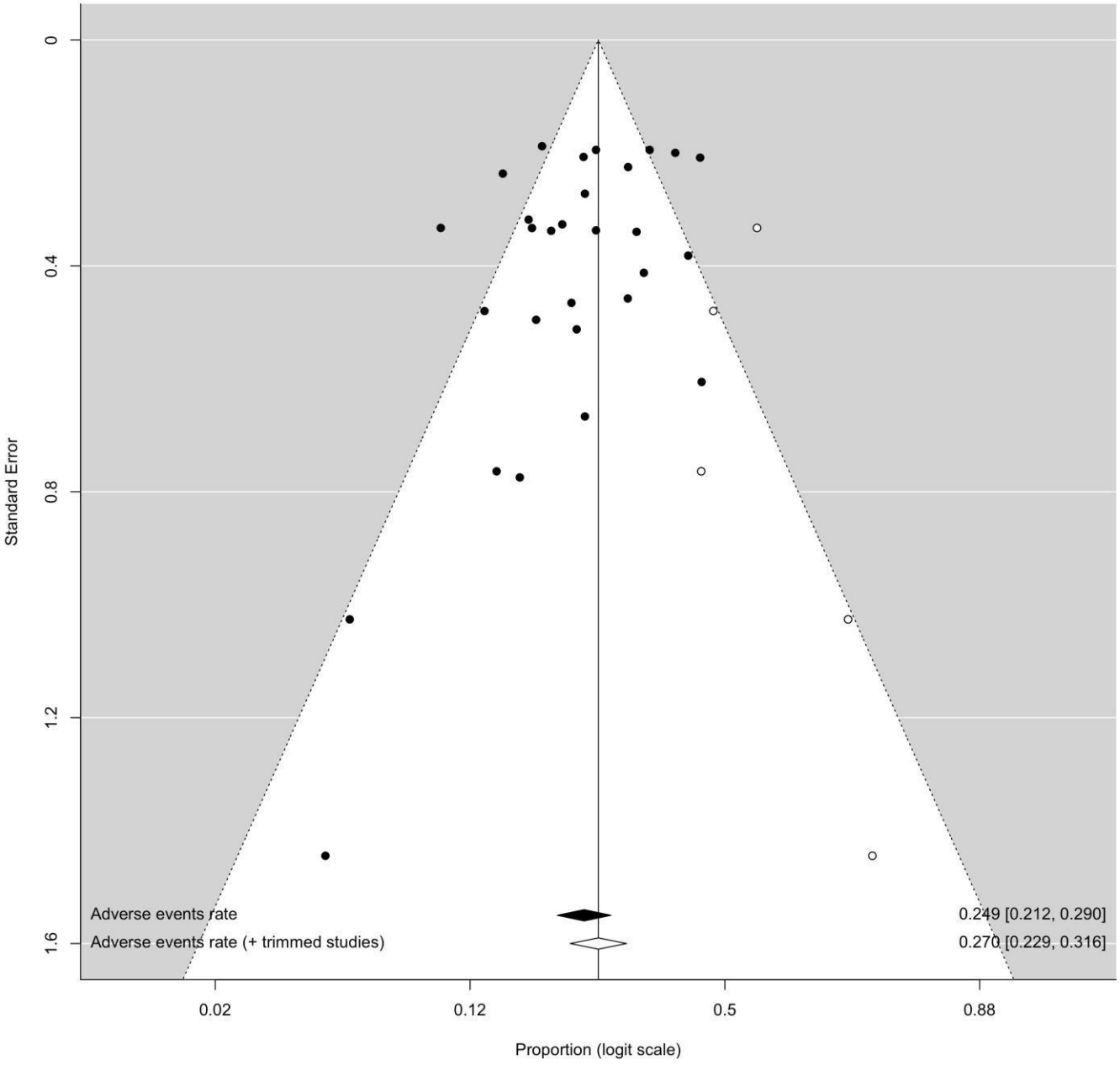
Supplementary Figure 5: Forest plot reporting the rates of definitive treatment. CI: confidence interval. RE: random effect.



Supplementary Figure 6: Forest plot reporting the rates of lesions endoscopically managed. CI: confidence interval. RE: random effect.



Supplementary Figure 7: Funnel plot reporting the risk of publication bias.



Appendix 1

Data sources and search strategy

We performed a comprehensive literature search by using PubMed, EMBASE and SCOPUS (up to September 31$_{st}$ 2018) to identify full articles evaluating outcomes of endoscopic papillectomy for the treatment of ampullary lesions. PROSPERO was searched for ongoing or recently completed systematic reviews. Electronic searches were supplemented by manual searches of references of included studies and review articles.

We identified studies using the following medical subject headings (MeSH) and keywords including: “endoscopic papillectomy” and “endoscopic ampullectomy”. The search was restricted to English language.

The Medline search strategy was: "endoscopy"[MeSH Terms] OR "endoscopy"[All Fields] OR "endoscopic"[All Fields]) AND papillectomy[All Fields] OR ("endoscopy"[MeSH Terms] OR "endoscopy"[All Fields] OR "endoscopic"[All Fields]) AND ampullectomy[All Fields].

Selection process

Two review authors (M.S.; L.F.) independently screened the titles and abstracts yielded by the search against the inclusion criteria. Full reports were obtained for all titles that appeared to meet the inclusion criteria or where there was any uncertainty. Review author pairs then screened the full text and abstract reports and decided whether these met the inclusion criteria. Disagreements were resolved through discussion of all the authors. The reasons for excluding trials were recorded. Neither of the review authors was blinded to the journal titles or to the study authors or institutions. When there were multiple articles for a single study, we used the latest publication and supplemented it, if necessary, with data from the more complete version.

Inclusion and exclusion criteria

For the purpose of this systematic review, we included all clinical studies enrolling at least 10 patients with ampullary lesions treated by EP and reporting the rate of adverse events. Prospective and retrospective studies, published as full text, were considered. Studies not explicitly stating the endoscopic resectability criteria assessed during the pre-operative work up (i.e. dimensional limit, absence of intraductal ingrowth or endoscopic signs of malignancies or deep invasion) were excluded. Studies not published in the English language were excluded.

Data extraction

Using standardized forms, two reviewers (M.S., L.F.) extracted data independently and in duplicate from each eligible study. Reviewers resolved disagreements by discussion. Unresolved disagreements were resolved by two arbitrators (L.F., A.R.).The following data were extracted for each study including the publication status, the study design and location, the number of centers involved, the enrolment period, the number of all lesions treated and the number of sporadic or familiar lesions, the mean number of lesions per center per year, patient characteristics (average age, gender), mean tumor size, clinical presentation (Jaundice, Pain, Cholangitis, Pancreatitis, Asymptomatic), intraductal ingrowth, preoperative work-up, prophylactic biliary stenting and/or pancreatic stenting, number of lesions completely resected, number of lesions resected en-bloc, number of lesions with curative resection, need for same-session adjunctive treatments, adverse events (bleedings, perforations, pancreatitis, cholangitis papillary stenosis, deaths) and need for tratments (blood transfusions, embolizations, surgery) histology of the resected lesions (ie, adenoma, submucosal cancer and deep of infiltration, other histologies), number of patients in follow-up, mean follow-up period, number of patients

with recurrence, endoscopic re-treatment, incompleteness related surgery, recurrence related surgery and overall need for surgery, number of patient with curative resection and no recurrences, number of patients managed only by endoscopy.

Quality assessment

Quality was assessed by the modified Newcastle-Ottawa Scale for non-randomized studies, ranging from 0 (low-quality) to 5 (high-quality). Two reviewers (MS, LF) assessed quality measures for included studies and discrepancies were adjudicated by collegial discussion.

Outcomes assessment

The primary outcome was the rate of adverse events on a per-lesion basis. Secondary outcomes were the rate of complete resection, the rate of en-bloc resection, the rate of curative resection and the rate of cases needing a same-session adjunctive treatment if provided. The rate of recurrence, on a per-lesions basis, and both the rate of recurrence endoscopically re-treated and the rate of those surgically treated, were also assessed if provided. Further secondary outcomes were the prevalence of adenomas and the prevalence of invasive cancer. Finally the rate of definitive resection and the rate of endoscopically-managed lesions were also reported.

Outcomes Definitions

Adverse events: bleeding, perforation, post-papillectomy pancreatitis, cholangitis and papillary stricture were regarded as adverse events. Any intra- or post-procedural adverse events were reported.

Complete endoscopic resection: endoscopic resection was classified as "Complete" in the absence of any adenomatous remnant from the resection margins at the end the procedure, after meticulous endoscopic inspection. This indicator of technical success was not considered to be affected by any same-session adjunctive endoscopic treatment (i.e., soft tip coagulation or Argon Plasma Coagulation of margins).

Curative resection: endoscopic resection was considered as "Curative" in the absence of any histological features predicting a loco-regional persistence of dysplastic/neoplastic disease:

- Lateral or vertical margins testing positive for the excised lesion.
- Evidence of submucosal cancer or deeper infiltration.

Definitive resection: endoscopic resection was regarded as "Definitive" in the absence of any recurrence in the follow-up period after obtaining a Curative resection.

Endoscopically managed lesion: ampullary lesions were considered as "Endoscopically managed" in the presence of a complete excision of the lesion, regardless of the number of sessions required and the detection of a recurrence in the follow-up period, if this was amenable to endoscopic treatment again.

Statistical Analysis

The pooled proportions and rates were calculated using a random effects model in case of substantial heterogeneity across studies; otherwise, we used a fixed effect model[1]. Dependent variables were modeled on the logit (log-odds) scale and then converted to percentages with 95% confidence intervals (95%CIs). We computed statistical heterogeneity with the I^2 statistic (high heterogeneity level >50%) and tested it using the Q2 test (statistical significance level set as $p< 0.1$). We developed subgroup meta-analytic models selecting a priori variables which could affect the outcomes, such as the mean lesion size. We also explored possible relationship between study variables and outcomes by means of univariate

metaregression.

Publication bias was assessed by means of funnel plot with trim-and-fill analysis and by Begg and Mazumdar's test[2].

All the analyses were performed using R statistical software with metafor package. All tests were 2-sided, and statistical significance was considered as $p< 0.05$, excluding the investigation of heterogeneity across studies in which $p< 0.10$ was regarded as significant.

References

1. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. Contemporary clinical trials 2015;45:139-45.
2. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088-101.

Appendix 2

Study characteristics and quality

All studies were published between 2002 and 2018. The study countries were the following: 10 were performed in Asia (468 lesions), 9 in the United States (774 lesions), 9 in Europe (384 lesions), and 1 in Australia (125 lesions). Six studies had a prospective design and the other 23 were retrospective. Twenty-five of 29 were single-center studies, otherwise 4 studies reported results of multicenter experiences. Study characteristics are briefly reported in Appendix 2, Supplementary Table8.

In total, the 29 studies included 1751 ampullary lesions which were removed endoscopically. The 66.8% (52.5-81.0%) of the lesions were sporadic ampullary lesions. The remaining lesions were related to Familial Adenomatous Polyposis. Patients had a mean age of 60.2 (56.9-63.5) years, most of them being male (617/1178 patients, 24 studies). Among the 16 studies reporting data on clinical symptoms leading to the endoscopic treatment, 485 of 1081 ampullary lesions were asymptomatic. On the other hand the 14.4% (6.9-21.9%), 16.6% (7.1-26.0%), 4.1% (1.9-6.3%), and 1% (0.3-1.6%) of patients were reported to suffer from abdominal pain, jaundice, pancreatitis and cholangitis, respectively.

The mean lesion size was 15.7 mm (13.1-18.3;). All the studies reported histological features of resected ampullary lesions: the 80.4% (74.6-86.3%) of lesions were adenomas, with the 19.7% (13.4-26.0%) of them having high grade dysplasia (HGD) stigmata. Otherwise, the 11.3% (7.7-14.8%) were found to be malignant adenocarcinomas. The average Newcastle Ottawa score of the studies was 5.7 (range 4-6).

Subgroup analysis on placing prophylactic pancreatic stenting or not.

A prophylactic pancreatic stent was placed after most of the resections (77.4% of cases, 66.7-85.4) across 21 studies (1262 lesions). Pancreatic stent characteristics were reported in appendix 2, Table 10.

Twelve of the included studies provided data on post-procedural pancreatitis occurring after prophylactic pancreatic stent placement; the calculated pooled rate was 6.9% (4.8-10.0; I^2=0%).

Otherwise the post-procedural pancreatitis rate was 20.7% (13.5-30.4; I^2=0%) among patients in whom a prophylactic pancreatic stent was not placed. Prophylactic pancreatic stent placement significantly decrease the risk of post procedural pancreatitis (p: < 0.001).

Follow-up:

Mean post resection follow up, ranging from 9.6 to 84.5 months, was reported by 25 studies including 1331 ampullary lesions. The pooled recurrence rate was 11.8% (8.3-16.5; I^2=72%). (Supplementary Figure 4).

The pooled rate of ampullary lesions definitively managed through single endoscopic approach was 72.5% (64.5-69.3) (Supplementary Figure 5).

On the other hand, 7.5% (4.6-11.9; I^2=73%) of patients benefited from an endoscopic re-treatment. Thus 80.9% (72.7-87.0) of ampullary lesions effectively treated by endoscopic fashion only (Supplementary Figure 6). An overview of the main follow-up outcomes is reported in Supplementary Table9.

Based on the data reported by 24 studies, 95 patients were referred for surgery following a non-curative endoscopic resection, yielding a pooled rate of 7.7% (4.4-13.2; I^2=82%). Moreover, surgery was needed in the 4.1% (2.8-6.2; I^2=32%) of patients due to recurrent lesions.

An overall pooled rate of 9.7% (6.0-15.3; I^2= 83%) patients underwent surgery across the included studies.

Publication bias on primary outcome (overall adverse events).

Significant publication bias was found according to Egger's test for funnel plot asymmetry (z = -2.0796, p = 0.0376). On trim-and-fill analysis, 5 missing studies were found, all on the right side of the pooled estimate, i.e. yielding a higher adverse events rate. However, adding trimmed studies did not significantly affect the outcome estimate (Supplementary Figure 7), therefore publication bias was judged as trivial.

Supplementary table 1. Univariate metaregression on overall adverse events rate. Analysis is conducted on logarithmic scale.

Variable	Coefficient	Standard error	P value	R^2 (amount of heterogeneity accounted for)
Mean age	0.01 (-0.03 to 0.04)	0.02	0.748	0%
Mean number of lesions per center per year	0.01 (-0.02 to 0.04)	0.02	0.421	0%
Mean tumor size	-0.02 (-0.10 to 0.06)	0.04	0.561	0%
Asymptomatic onset (proportion)	0.69 (-0.19 to 1.56)	0.44	0.123	10%
En bloc resection rate (proportion)	0.93 (-0.62 to 2.47)	0.79	0.241	12%
Biliary stenting (proportion)	-0.71 (-2.43 to 1.01)	0.88	0.420	0%
Pancreatic stenting (proportion)	-0.55 (-1.79 to 0.69)	0.63	0.386	7%
Adenocarcinoma (proportion)	1.31 (-0.39 to 3.02)	0.87	0.130	11%

Supplementary Table 2. Univariate metaregression on pancreatitis rate. Analysis is conducted on logarithmic scale.

Variable	Coefficient	Standard error	P value	R^2 (amount of heterogeneity accounted for)
Mean age	-0.01 (-0.04 to 0.02)	0.02	0.376	0%
Mean number of lesions per center per year	0.01 (-0.02 to 0.04)	0.02	0.334	8%
Mean tumor size	-0.03 (-0.01 to 0.05)	0.04	0.492	0%
Pancreatitis onset (proportion)	3.77 (-12.80 to 20.33)	8.45	0.656	0%
Pancreatic stenting (proportion)	**-1.72 (-2.95 to -0.50)**	**0.63**	**0.006***	**62%**
Adenocarcinoma (proportion)	1.44 (-0.48 to 3.37)	0.98	0.141	10%

Supplementary table 3. Overview of Adverse Events (AE). Causes of deaths were: necrotizing pancreatitis (3), retroperitoneal perforation (2), acute myocardial infarction (1) BT: blood transfusion.

Reference	Bleed	BT	Embolization	perforation	Pancreatitis	Severe pancreatitis	Cholangitis	Stenosis	Overall AEs	AEs surgery	Death
Norton 2002	2	0	0	1	4	0	0	2	9	0	0
Catalano 2004	2	0	0	0	5	0	0	3	10	1	0
Cheng 2004	4	\	\	1	5	0	0	2	12	0	0
Katsinelos 2006	1	0	0	0	1	0	0	0	2	0	0
Jung 2009	1	\	0	0	4	0	2	0	7	0	0
Ghidirim 2009	2			0	1	0	0	0	3	0	0
Boix 2009	1	\	0	0	4	0	0	0	5	0	0
Kim 2009	0	0	0	0	1	1	0	0	1	0	1
Irani 2009	5	\	1	2	10	0	1	3	21	1	0
Hwang 2010	4	0	0	0	0	0	0	1	5	0	0
Patel 2011	2	0	0	0	3	0	0	0	5	0	0
Ito 2012	7	0	0	2	0	0	3	0	12	0	0
Salmi 2012	3	\	\	2	6	1	0	0	11	1	0
Ceppa 2013	6	\	\	0	6	0	0	0	12	0	0
Kim 2013	12	\	0	0	6	0	0	0	18	0	0
Laleman 2013	11	3	1	0	14	1	4	0	29	0	0
Will 2013	3	\	\	1	7	0	0	0	11	0	0
Ridtitid 2014	23	\	\	3	9	0	0	0	35	1	1
Onkendi 2014	17	7	2	0	19	3	0	0	36	0	0
Napoleon 2014	11	5	0	4	22	4	5	0	42	3	1
Ma 2014	1	1	0		5	4	0	0	6	0	0
Shim 2014	4	0	0	2	7	2	0	0	13	0	0
De Palma 2015	2	0	0	0	3	0	0	0	5	0	0
Lee 2015	3	\	1	1	7	1	0	1	12	0	0
Tsuji 2015	21	9	1	3	12	1	0	5	41	0	1
Soma 2015	1	\	0	0	1	0	0	0	2	0	0
Dubois 2016	0	0	0	0	0	0	0	0	0	0	0
Kang 2017	18	\	0	8	16	0	0	0	42	4	2
Klein 2018	30	8	0	1	9	0	0	0	31	0	0

Supplementary table 4. Univariate metaregression on oncologically curative resection rate. Analysis is conducted on logarithmic scale.

Variable	Coefficient	Standard error	P value	R^2 (amount of heterogeneity accounted for)
Mean tumor size	-0.03 (-0.20 to 0.13)	0.08	0.715	0%
Mean number of lesions per center per year	0.02 (-0.03 to 0.06)	0.02	0.496	2%
En bloc resection rate (proportion)	3.55 (1.11 to 5.99)	1.24	0.004*	49%

Supplementary table 5. Overview of Resection Outcomes.

Reference	Complete resection	En Bloc	Curative resection	Need for Surgery (not curative)
Norton 2002	26	24	25	1
Catalano 2004	99	\	93	6
Cheng 2004	53	30	42	1 3
Katsinelos 2006	14	13	11	3
Jung 2009	17	17	12	6
Ghidirim 2009	11	8	10	2
Boix 2009	\	\	\	1 5
Kim 2009	20	\	18	1
Irani 2009	\	\	\	\
Hwang 2010	11	11	11	0
Patel 2011	38	36	38	0
Ito 2012	28	27	27	0
Salmi 2012	61	\	52	5
Ceppa 2013	\	\	68	0
Kim 2013	65	60	65	0
Laleman 2013	71	\	71	1 7
Will 2013	46	\	44	0
Ridtitid 2014	\	89	134	\
Onkendi 2014	126	120	126	\
Napoleon 2014	93	68	84	4
Ma 2014	26	17	26	0
Shim 2014	\	35	39	\
De Palma 2015	27	24	25	2
Lee 2015	45	/	38	1

Tsuji 2015	93	88	93	1
Soma 2015	12	12	12	\
Dubois 2016	8	\	7	4
Kang 2017	104	94	93	6
Klein 2018	122	NA	117	8

Supplementary table 6. Sensitivity analysis: patients and lesions characteristics.

Variable	Pooled estimates	I^2
Mean tumor size	16.8	98%
Mean age	63.0	50%
Onset		
Jaundice onset	19.5%	97%
Pain onset	14.4%	99%
Cholangitis onset	0.6%	1%
Pancreatitis onset	5.4%	95%
Histology		
Proportion of adenomas	78.6%	96%
Proportion of adenocarcinoma	10.4%	92%

Supplementary Table 7. Efficacy issues to be addressed by ideal future studies on endoscopic papillectomy.

HOW SHOULD IT BE?
-Prospective design -Multiple referral centers should be involved -Strict inclusion criteria based on: o Endoscopic features (EGD + ERCP) o Radiologic fetaures (MRI+MRCP or EUS) o Histologic features -Strict exclusion criteria: o FAP or less common familial conditions -Unique outcome definitions in terms of efficacy: o Technical success o En bloc resection o Complete resection (R0), histologically assessed as an en bloc resection with both lateral and deep margin free from dysplasia/neoplasia o Oncologically curative resection defined as an R0 resection without poor prognostic histologic features o Recurrence o Definitively curative resection, defined as cured lesion with no recurrences within an adequate follow up time. The endoscopic follow up should including biopsy sampling o Patients saved from surgery -Appropriate statistical analysis including cost/efficacy analysis and uni-multivariate analysis investigating factors predicting efficacy outcomes

Supplementary table 8: studies characteristics

Author	Year	Country	Design	Mono/Multicenter	Period	Lesions (n)
Norton	2002	USA	Retrospective	Mono	1997-1999	26
Catalano	2004	USA	Retrospective	Multi (4)	1998-2001	103
Cheng	2004	USA	Retrospective	Mono	1994-2003	55
Katsinelos	2006	Greece	Retrospective	Mono	1998-2004	14
Jung	2009	Korea	Retrospective	Mono	2003-2008	22
Ghidirim	2009	Moldova	Retrospective	Mono	1998-2008	12
Boix	2009	Spain	Retrospective	Mono	1995-2007	21
Kim	2009	Korea	Retrospective	Mono	1997-2008	20
Irani	2009	USA	Retrospective	Mono	1997-2007	141
Hwang	2010	Korea	Prospective	Mono	2007-2009	11
Patel	2011	USA	Retrospective	Mono	1996-2009	38
Ito	2012	Japan	Retrospective	Mono	2002-2010	28
Salmi	2012	France	Prospective	Mono	2002-2009	61
Ceppa	2013	USA	Retrospective	Mono	2000-2010	68
Kim	2013	Korea	Prospective	Mono	2005-2012	72
Laleman	2013	Belgium	Retrospective	Mono	2000-2008	91
Will	2013	Germany	Prospective	Mono	2005-2012	54
Ridtitid	2014	USA	Retrospective	Mono	1995-2012	182
Onkendi	2014	USA	Retrospective	Mono	1994-2009	135
Napoleon	2014	France	Prospective	Multi (12)	2003-2006	93
Ma	2014	USA	Retrospective	Multi (2)	2000-2010	26
Shim	2014	Korea	Retrospective	Mono	2006-2012	39
De Palma	2015	Italy	Retrospective	Mono	2008-2013	27
Lee	2015	Korea	Prospective	Mono	2012-2014	45
Tsuji	2015	Tokyo	Retrospective	Mono	1999-2014	115
Soma	2015	Tokyo	Retrospective	Mono	2008-2011	12
Dubois	2016	Switzerland	Retrospective	Mono	2005-2015	11
Kang	2017	Korea	Retrospective	Multi (5)	2007-2014	104
Klein	2018	Australia	Retrospective	Mono	2009-2017	125

Supplementary table 9. Overview of Follow-up (FU) Outcomes.

Reference	Patients in FU	FU (months)	Recurrences	Endoscopic re-tratment	Surgery (recurrence)	Definitive treatment	Endoscopically managed
Norton 2002	22	13	2	\	\	19	19
Catalano 2004	103	36	10	0	10	83	83
Cheng 2004	38	30	7	7	0	18	25
Katsinelos 2006	14	28,3	2	2	0	9	11
Jung 2009	18	9,6	1	0	1	7	7
Ghidirim 2009	10	\	0	0	0	10	10
Boix 2009	21	15,9	1	1	0	6	6
Kim 2009	20	20	4	2	2	14	16
Irani 2009	119	\	\	\	\	86	86
Hwang 2010	11	12,1	0	0	0	11	11
Patel 2011	38	17,2	6	6	0	32	32
Ito 2012	28	26,1	4	4	0	23	27
Salmi 2012	61	36	3	3	0	49	52
Ceppa 2013	\	\	\	\	\	\	\
Kim 2013	70	23,7	5	3	2	60	68
Laleman 2013	91	32	13	11	1	59	70
Will 2013	54	\	4	3	0	40	43
Ridtitid 2014	\	\	\	\	\	\	\
Onkendi 2014	135	53	44	\	\	\	\
Napoleon 2014	93	\	5	1	1	78	79
Ma 2014	24	84,5	14	14	0	10	24
Shim 2014	39	15	3	\	\	36	36
De Palma 2015	27	18,4	1	1	0	24	25
Lee 2015	45	13,5	0	1	0	38	44
Tsuji 2015	\	\	\	20	0	\	\
Soma 2015	12	28,5	3	\	\	8	\
Dubois 2016	\	\	\	1	\	\	8
Kang 2017	99	44,2	\	5	\	89	94
Klein 2018	113	18,5	18	0	\	88	\

Supplementary table 10. Pancreatic stents. All the authors of the included studies used plastic pancreatic stent.

Reference	Lenght (cm)	Diameter (Fr)	Producer
Norton 2002	\	5	\
Catalano 2004	\	\	\
Cheng 2004	\	3-5	Wilson-Cook Medical Inc
Katsinelos 2006	\	5	\
Jung 2009	\	\	\
Ghidirim 2009	4-5	4-7	Wilson-Cook Medical Inc
Boix 2009	\	\	\
Kim 2009	\	\	\
Irani 2009	8	3	\
Hwang 2010	\	5	Daikin Industries
Patel 2011	\	3-5	\
Ito 2012	\	5	\
Salmi 2012	3	5	Wilson-Cook Medical Inc
Ceppa 2013	\	\	\
Kim 2013	\	5	Wilson-Cook Medical Inc
Laleman 2013	\	\	\
Will 2013	\	5	GIP Medizintechnik GmbH
Ridtitid 2014	\	3-5	\
Onkendi 2014	\	\	\
Napoleon 2014	\	\	\
Ma 2014	\	\	\
Shim 2014	\	\	\
De Palma 2015	\	5-7	\
Lee 2015	5-7	5	Wilson-Cook Medical Inc
Tsuji 2015	5-7	5	\
Soma 2015	\	\	\
Dubois 2016	\	\	\
Kang 2017	\	\	\
Klein 2018	\	\	\