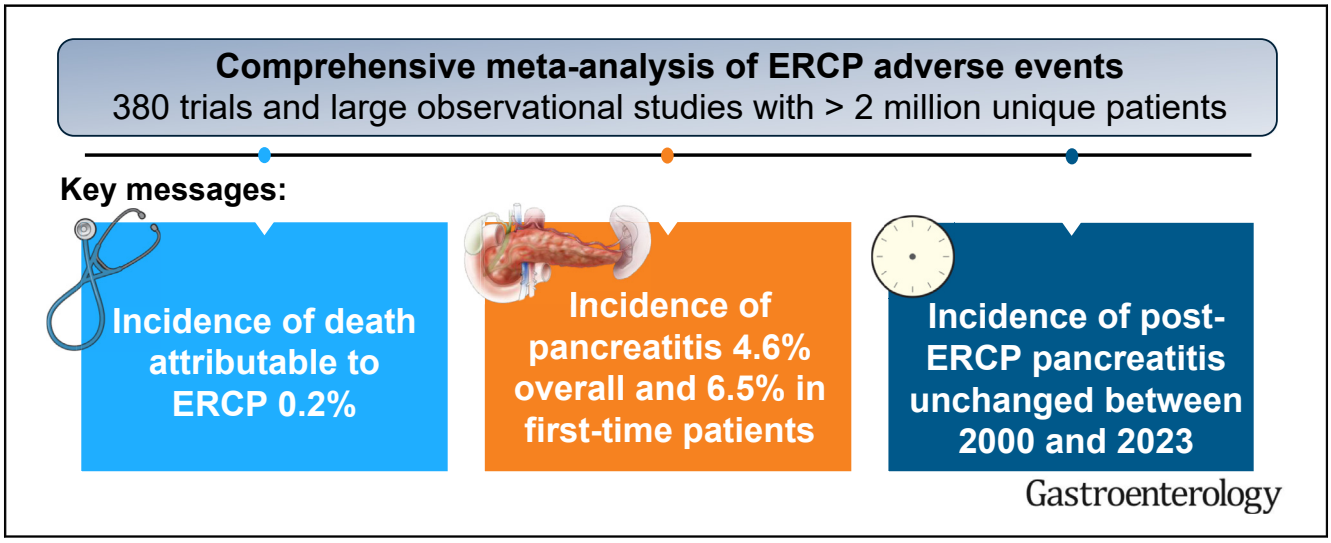
Gastroenterology 2025;168:568–586

Adverse Events Associated With Endoscopic Retrograde Cholangiopancreatography: Systematic Review and Meta-Analysis

Kirles Bishay,[1,2,*] **Zhao Wu Meng**,[1,2,*] Rishad Khan,[3] Mehul Gupta,[4] Yibing Ruan,[5,6] Marcus Vaska,[7] Jordan Iannuzzi,[1] Dylan E. O'Sullivan,[6] Brittany Mah,[5,6] Arun C. R. Partridge,[4] Amanda M. Henderson,[8] Howard Guo,[1] Sunil Samnani,[4] Max DeMarco,[1] Yuhong Yuan,[9,10] B. Joseph Elmunzer,[11] Rajesh N. Keswani,[12] Sachin Wani,[13] Zachary L. Smith,[14] Ronald J. Bridges,[1] Steven J. Heitman,[1,2] Robert J. Hilsden,[1,2] Darren R. Brenner,[2,5,6] Grigorios I. Leontiadis,[10] and Nauzer Forbes[1,2]

[1]Division of Gastroenterology and Hepatology, Department of Medicine, University of Calgary, Calgary, Alberta, Canada; [2]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada; [3]Division of Gastroenterology and Hepatology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [4]Department of Medicine, University of Calgary, Calgary, Alberta, Canada; [5]Department of Oncology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [6]Department of Cancer Epidemiology and Prevention Research, Cancer Control Alberta, Alberta Health Services, Calgary, Alberta, Canada; [7]Knowledge Resource Service, Alberta Health Services, Calgary, Alberta, Canada; [8]Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [9]Department of Medicine, London Health Science Centre, London, Ontario, Canada; [10]Farncombe Family Digestive Health Research Institute, Department of Medicine, McMaster University, Hamilton, Ontario, Canada; [11]Division of Gastroenterology and Hepatology, Medical University of South Carolina, Charleston, South Carolina; [12]Division of Gastroenterology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois; [13]Division of Gastroenterology and Hepatology, University of Colorado Anschutz Medical Campus, Aurora, Colorado; and [14]Division of Gastroenterology and Hepatology, Medical College of Wisconsin, Milwaukee, Wisconsin



BACKGROUND & AIMS: Endoscopic retrograde cholangiopancreatography (ERCP)–related adverse events (AEs) are associated with morbidity, mortality, and health care expenditure. We aimed to assess incidences and comparisons of ERCP AEs. **METHODS:** We included studies performed after 2000 reporting on ERCP AEs from database inception through March 12, 2024. Outcomes included pancreatitis, bleeding, cholangitis, cholecystitis, perforation, and death. DerSimonian and Laird random effects meta-analyses were performed to calculate incidences of AEs. Subgroup and pairwise meta-analyses were performed. Meta-regression was performed on median recruitment year to assess temporal trends in pancreatitis incidence. **RESULTS:** A total of 380 studies were included. The incidence of death attributable to ERCP was 0.2% (95% confidence interval [CI], 0.1%–0.3%; I^2, 44%; n = 47,258) in all-comers. The overall incidence of pancreatitis was 4.6% (95% CI, 4.0%–5.1%; I^2, 96%; n = 293,378) among all-comers and 6.5% (95% CI, 5.9%–7.1%, I^2, 89%; n = 88,809) among first-time patients. Pancreatitis incidence remained stable between 2000 and 2023 (average annual percent change 0.06, 95% CI, −0.27 to 0.39). The overall incidences of the following AEs for all-comers were: bleeding (1.5%; 95% CI, 1.2%–1.7%;

I^2, 93%; n = 229,655), cholangitis (2.5%; 95% CI, 1.9%–3.3%; I^2, 96%; n = 121,619), cholecystitis (0.8%; 95% CI, 0.5%–1.2%; I^2, 39%; n = 7799), and perforation (0.5%; 95% CI, 0.4%–0.6%; I^2, 90%; n = 306,378). **CONCLUSIONS:** ERCP-associated AEs remain common. Incidence of post-ERCP pancreatitis remained static despite improvements in techniques, prevention, and recognition. These results are important to patients, endoscopists, and policy makers to inform consent and to encourage implementation of available risk mitigation strategies.

Keywords: ERCP; Pancreatitis; Bleeding; Adverse Event; Complication; Quality.

Endoscopic retrograde cholangiopancreatography (ERCP) is commonly performed for pancreaticobiliary conditions and is technically complex.[1–3] Unsurprisingly, ERCP is associated with a high rate of adverse events (AEs).[4,5] ERCP procedural volumes in the United States have remained stable over time[6,7] despite the advent of alternate endoscopic strategies with lower AE rates,[8] underscoring not only the continued importance of ERCP but also pointing to potential overuse.[9] ERCP is currently performed by both low- and high-volume providers and at low- and high-volume centers, with significant variability in reported AEs.[10,11]

ERCP-related AEs include post-ERCP pancreatitis (PEP), bleeding, cholangitis, cholecystitis, perforation, and even death. The collective incidence of ERCP-associated AEs has been demonstrated to exceed 10%,[4,5,10,11] with PEP being the most well-studied. Despite concerted and widespread efforts to improve ERCP quality, identify risk factors for PEP, and implement protective measures, the mortality associated with PEP is rising.[12,13] The average costs associated with PEP exceed $10,000 USD per admission,[14,15] suggesting that PEP accounts for more than $300 million in annual US health care spending.[16–18]

Despite an abundance of primary literature describing ERCP-associated AEs, there remains a relative lack of high-quality evidence syntheses. A recent study reported the pooled incidence of PEP from more than 140 randomized controlled trials (RCTs).[12] Importantly, this study did not calculate incidences of any other AEs. In addition, no evidence syntheses have reported incidences of PEP from observational studies, which assess populations that are often more generalizable to "real-world" clinical settings and are thus necessary to fully understand ERCP AEs.[19,20] Furthermore, no high-quality meta-analyses have been performed of AE rates in specific clinically relevant subgroups, including those established as risk factors for AEs.[21]

Given that ERCP-related AEs are common and are associated with significant morbidity, mortality, and health care burden, high-quality contemporary pooled estimates and comparisons are needed to inform patients, providers, and other stakeholders of procedural risks. We aimed to bridge these important gaps by conducting a comprehensive contemporary systematic review and meta-analysis of ERCP AEs from both RCTs and observational studies.

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Endoscopic retrograde cholangiopancreatography is a commonly performed endoscopic procedure with a relatively high rate of adverse events.

NEW FINDINGS

In this meta-analysis of 380 studies, death attributable to endoscopic retrograde cholangiopancreatography was 0.2% in all-comers. Pancreatitis incidence remained stable between 2000 and 2023. The incidences of the adverse events for all-comers were as follows: bleeding (1.5%), cholangitis (2.5%), cholecystitis (0.8%), and perforation (0.2%).

LIMITATIONS

There was substantial to considerable statistical heterogeneity in several of our pooled estimates. Despite meta-regression analyses for patient sex, patient age, and trainee involvement, there still remained substantial heterogeneity.

CLINICAL RESEARCH RELEVANCE

Our study provides contemporary data informing estimated incidences of all common endoscopic retrograde cholangiopancreatography–related adverse events and estimates of magnitudes of risk associated with clinically relevant patient- and procedure-related factors, raising awareness and facilitating implementation of evidence-based interventions to mitigate risks.

BASIC RESEARCH RELEVANCE

The incidence of post–endoscopic retrograde cholangiopancreatography pancreatitis has remained static over the past generation. The study of novel pharmacologic interventions and other basic approaches to mitigate this risk should be a priority.

Methods

Overview and Rationale

We conducted this meta-analysis according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Meta-analysis Of Observational Studies in Epidemiology (MOOSE) recommendations[22,23] (checklists in Supplementary Tables 1 and 2). Our protocol was registered *a priori* (CRD42020220221) and published.[24] Research ethics board approval was not required because we used publicly available data.

We aimed to produce contemporary accurate estimates of incidences of ERCP-related AEs from RCTs and observational

*** Authors share co-first authorship.**

Abbreviations used in this paper: **AE, adverse event; CI, confidence interval; ERCP, endoscopic retrograde cholangiopancreatography; PEP, post-ERCP pancreatitis; RCT, randomized controlled trial; RR, risk ratio.**

Most current article

0016-5085
https://doi.org/10.1053/j.gastro.2024.10.033

studies, both separately and pooled together, and to report AE incidences within relevant subgroups in addition to performing pairwise comparisons to offer relative estimates of risk. Our rationale to report the results from RCTs and observational studies separately stems from potentially important differences in underlying populations included,[25] outcomes ascertainment approaches,[26] and event reporting,[27] among other considerations. Within the context of this review, however, for the purposes of yielding data on incidences of AEs (rather than comparisons of effectiveness of interventions, for example), it could be reasonably argued that RCT data are ultimately no different from those from observational studies. Therefore, the decision was made to report all results from RCTs and large observational studies both separately *and* pooled together.

Data Sources and Searches

We carried out 2 electronic searches, 1 for RCTs and 1 for observational studies, both with the oversight of an experienced health research librarian (Marcus Vaska). Searches were performed of MEDLINE (Ovid), PubMed, CINAHL, EMBASE, Scopus, Web of Science, and Evidence-Based Medicine Reviews. The initial title and abstract search was performed from inception to November 10, 2020, updated on November 21, 2022, then re-updated on March 12, 2024. We used Medical Subject Heading (MeSH) and free-text terms with spelling variations and synonyms (Supplementary Table 3).

Study Selection

An RCT was included if it met *all* the following criteria: (1) it presented original data with any primary research question; (2) either the intervention or control arm comprised adult patients undergoing ERCP for any indication(s); (3) it reported the incidence of an ERCP-related AE (1 or more of PEP, bleeding, cholangitis, cholecystitis, perforation, or death); and (4) $\geq$75% of included patients underwent ERCP in the year 2000 or later, with this cutoff chosen to include only studies representative of the current "era" of ERCP. An RCT was excluded if *any* of the following criteria were met: (1) it was not written in English; (2) it was a conference abstract; or (3) the data contained therein overlapped with data from another included study, in part or in whole. In this case, the study with more patients was included.

An observational study was included if it met *all* the following criteria: (1) it presented original data with any primary research question; (2) the primary or secondary objective of the study was to report the incidence of 1 or more ERCP-related AE(s); and (3) $\geq$75% of included patients underwent ERCP in the year 2000 or later. An observational study was excluded if *any* of the following criteria were met: (1) it was not written in English; (2) it was a conference abstract; (3) the data contained therein overlapped with data from another included study, in part or in whole; (4) it represented the experience of a single endoscopist; or (5) there were fewer than 500 patients included. The decision to exclude studies with fewer than 500 patients was made to mitigate small-study effects and to reduce the likelihood of including zero-event studies.

We imported all citations into Covidence (Melbourne, Australia) and removed duplicates. Nine reviewers (Kirles Bishay, Zhao Wu Meng, Jordan Iannuzzi, Dylan E. O'Sullivan, Brittany Mah, Arun C.R. Partridge, Amanda M. Henderson, Max DeMarco, Nauzer Forbes) were randomly assigned roughly equal numbers of citations at each review stage. Each record was screened independently by 2 reviewers in duplicate, with a vote of "both include" resulting in the record proceeding to the full-text phase, and with discrepancies resolved by the senior author (Nauzer Forbes). Included records underwent full-text screening by 2 of the preceding reviewers in duplicate, with discrepancies resolved by consensus.

Data Extraction and Risk of Bias Assessment

We extracted data into standardized forms that included study details, patient and endoscopist data, AEs, and outcome definitions, including according to prespecified subgroups (see later in this article) and across relevant pairwise comparisons. For RCTs, data were extracted using the intervention ultimately received, including if crossed over from a different randomly assigned intervention. Outcomes included PEP, bleeding, cholangitis, cholecystitis, perforation, and procedure-related death. The most commonly used definitions of ERCP-related AEs are the consensus definitions,[28] which have since been further formalized within the American Society for Gastrointestinal Endoscopy Lexicon[29] and used as the basis for more complex schemas and causal attribution systems within ERCP.[30] This set of definitions was therefore used as the standard for all ERCP-related AEs. Given the associated heterogeneity in outcomes and definitions, neither of sedation-related adverse events nor unplanned health care presentations for non-AE reasons were analyzed. Two authors (Howard Guo and Sunil Samnani) conducted risk of bias assessments in duplicate with the Cochrane Risk of Bias tool, version 2 (RoB 2)[31] for RCTs and the Newcastle-Ottawa Scale[32] for observational studies. Discrepancies were resolved by consensus.

Data Synthesis and Analysis

We used DerSimonian and Laird random effects meta-analyses to report incidences of individual post-ERCP AEs with 95% confidence intervals (CIs) from RCTs and observational studies (1) separately and (2) pooled together. For all estimates of AE incidences, both overall and within subgroups, data from RCTs were treated the same as those from observational studies (ie, both were considered "observational-type" data) given that the randomization element was not relevant. Similarly, for most pairwise comparisons of relative risks, RCT data were treated as observational if the variables in question were not used for randomization (eg, female vs male sex, or inadvertent pancreatic duct cannulation vs none). However, for certain pairwise comparisons (eg, use of nonsteroidal anti-inflammatory drugs vs placebo, or use of double guidewire technique vs pre-cut sphincterotomy), the interventional arm to which patients were randomized was used for pooled group assignment. In cases in which observational-type data from RCTs were pooled with data from observational studies, *P* values were provided for comparisons between the individual groups. We measured study weights using the inverse variance method. We measured statistical heterogeneity using the I^2 statistic and considered an I^2 of >50% as indicative of substantial heterogeneity. We assessed publication bias using Egger's test[33] and through visual inspection of funnel plots. Proportional meta-analyses were performed to assess the incidences of mild, moderate, and severe PEP per 100 incident cases in addition

to PEP-related death. All analyses were conducted using R (R Foundation for Statistical Computing, Vienna, Austria).

All outcomes were calculated for 2 nonoverlapping cohorts: (1) only first-time ERCP patients (ie, those with native papillary anatomy) and (2) all-comers to ERCP (ie, cohorts representative of standard practice, containing first-time patients but also those undergoing repeat ERCPs). "High-risk" patients (ie, those specifically at higher risk of PEP based on patient- and/or procedure-related eligibility criteria) were analyzed separately from the preceding 2 cohorts and presented as their own subgroup. High-risk patients were eligible for inclusion in relevant pairwise analyses but were excluded from any subgroup or sensitivity analyses given that they would falsely elevate AE incidence estimates.

We planned to perform several subgroup analyses based on a priori selected patient factors, procedural indications, intraprocedural factors and techniques, and techniques to mitigate PEP.[24] Data informing incidences for relevant procedural parameter-related subgroups were only extracted from studies that reported data from a complementary subgroup. For instance, data on patients undergoing biliary sphincterotomy were taken only from studies that also had data available on balloon dilation alone or sphincterotomy with balloon dilation. This was done to avoid including data within inappropriate contexts—for instance, using the preceding example, if data on patients with sphincterotomy were extracted from a study only reporting this value, it would be unknown whether "sphincterotomy" referred to sphincterotomy alone or with balloon dilation. Subgroup analyses of pooled AE incidences were performed if there were 3 or more studies with available data.

Pairwise meta-analyses were also performed for any comparisons with 3 or more studies available to inform an analysis. For all pairwise analyses, we assessed the certainty of the evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework.[34] As described previously, data on comparisons of interventions from RCTs not randomizing patients according to those interventions were treated as "observational-type data" from randomized studies and correspondingly treated with a lower starting "certainty of evidence" grade. For comparisons of interventions aiming to mitigate PEP, data were only included if patients were randomized according to the intervention. For each comparison, we assessed inconsistency, indirectness, imprecision, and other sources of bias to yield overall estimates of the certainty of the evidence base. For all pairwise analyses, risk ratios (RRs) were reported along with corresponding 95% CIs.

Sensitivity analyses were also performed according to relevant methodological considerations, including primary study location (North American vs European vs Asian-Pacific), median patient recruitment dates (2000–2007 vs 2008–2014 vs 2015–2023), study design (single center vs multicenter), and outcome definitions used (Lexicon definition vs "other," which included lack of definition, ambiguity regarding definitions, or alternate definitions). Univariable meta-regression using sex, trainee involvement, and mean age was performed to explore heterogeneity in estimates of AE incidences between studies in those with available data. Finally, meta-regression was also performed on median recruitment year, both as a linear term and with restricted cubic spline (with 3 knots) to determine whether temporal trends existed in PEP incidence and whether there was a potentially nonlinear relationship. This analysis was performed for all-comers, in first-time patients, and all combined patients, with studies specifically recruiting high-risk patients excluded as described previously.

Results

Overview and Descriptive Results

We identified 15,949 records and performed full-text review of 658 records. Of these, we included 136 observational studies and 244 RCTs in the meta-analysis. A PRISMA flow diagram outlining the selection process is provided in Figure 1. Characteristics of included RCTs and observational studies are summarized in Supplementary Tables 4 and 5, respectively.

A total of 70,773 patients were included from RCTs. Data from 53,550 patients were used to inform incidence estimates for PEP from 190 RCTs—13,566 all-comers to ERCP and 39,984 first-time patients. From 109 observational studies, data informing PEP incidence were used from 1,502,279 all-comers and 48,825 first-time patients.

Overall Incidences of ERCP-associated AEs

The incidence of PEP from RCTs was 5.5% (95% CI, 4.5%–6.7%) in all-comers and 7.0% (95% CI, 6.3%–7.7%) among first-time patients, with I^2 of 79% and 80%, respectively. From observational studies, the corresponding PEP incidences in all-comers and first-time patients were lower, at 4.1% (95% CI, 3.5%–4.7%) and 5.2% (95% CI, 4.2%–6.4%), respectively (I^2 98% and 94%), with P values of $<.01$ for both comparisons. These results and overall incidences of bleeding, cholangitis, cholecystitis, perforation, and death attributable to ERCP are provided in Table 1 for all-comers and first-time patients from RCTs and large observational studies. The incidence of death attributable to ERCP was 0.5% (95% CI, 0.3%–0.8%) for all-comers and 0.3% (95% CI, 0.3%–0.4%) for first-time patients using RCT data with protocolized follow-up and event adjudication protocols, whereas the according rates for all-comers and first-time patients from large observational studies were 0.1% (0.1%–0.2%) and 0.1% (0.1%–0.2%), respectively ($P < .01$ for both comparisons). Forest plots of the pooled incidences of death attributable to ERCP in all-comers are provided in Figure 2. The estimated incidences of clinically significant bleeding ranged between 1.4% and 1.9%, of cholangitis between 1.5% and 4.5%, of cholecystitis between 0.5% and 1.5%, and of perforation between 0.4% and 0.8%. Individual forest plots for these incidences are provided in Supplementary Figures 1 to 24.

Severity of PEP and PEP-related Death

From proportional meta-analyses of 128 RCTs that each necessarily reported separate incidences of mild, moderate, and severe PEP, these proportions were 65.9% (95% CI, 62.0–69.6%) for mild PEP, 27.9% (95% CI, 24.7%–31.3%) for moderate PEP, and 10.2% (95% CI, 8.7%–11.8%) for severe PEP, with forest plots presented in Supplementary

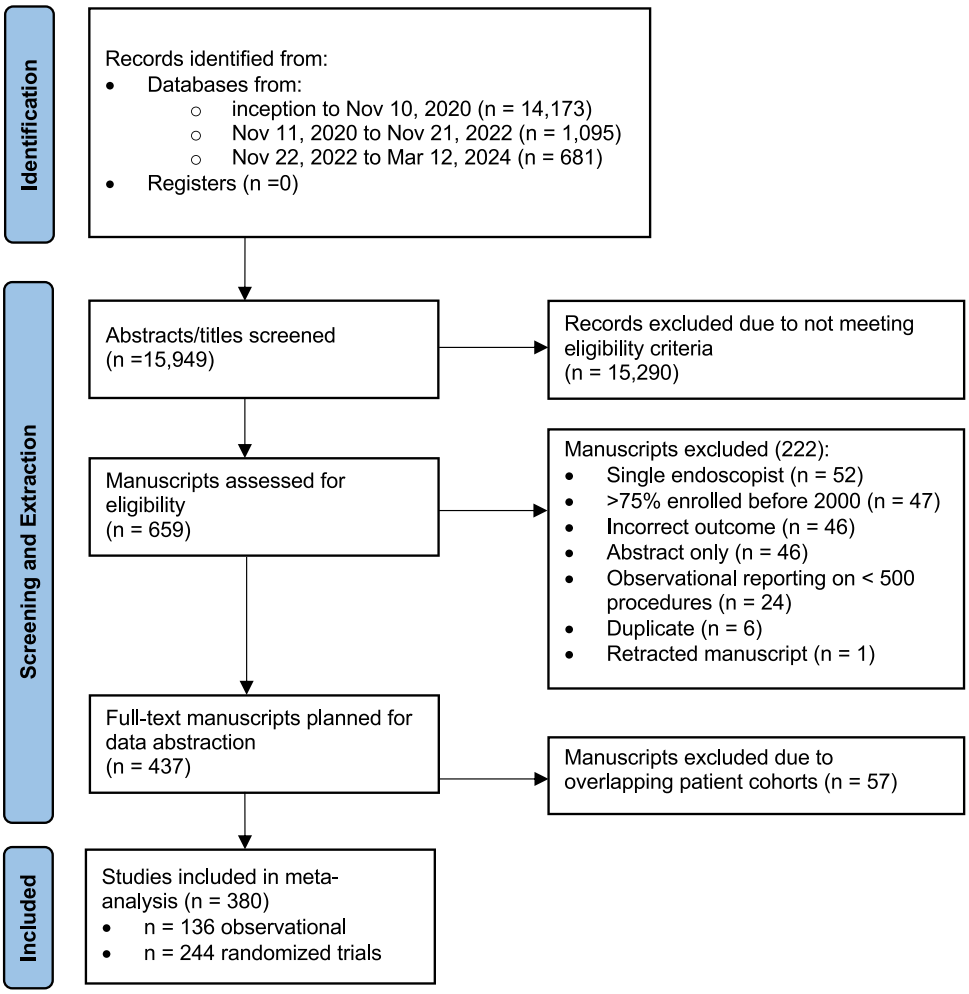


Figure 1. PRISMA study flow diagram summarizing study selection process.

Figures 25 to 27. Death resulted from PEP in 2.6% (95% CI, 1.8%–3.9%) of cases (Supplementary Figure 28).

Subgroup Analyses

Estimates of AE incidences within relevant subgroups and across relevant methodological considerations are summarized in Table 2, with forest plots provided in Supplementary Figures 29 to 85. Subgroups with the highest incidence of PEP based on pre-procedural factors included female patients (7.5%; 95% CI, 6.4%–8.6%), patients with confirmed or suspected sphincter of Oddi dysfunction (15.9%; 95% CI, 12.1%–19.1%) or patients with planned pancreatic interventions as an indication (12.8%; 95% CI, 6.3%–24.2%). In terms of procedural parameters, difficult cannulation resulted in a high incidence of PEP (11.4%; 95% CI, 9.2%–13.9%), with study-specific definitions of difficult cannulation and of high-risk patients overall provided in Supplementary Tables 6 and 7, respectively. The performance of a pre-cut (transpancreatic or suprapapillary) sphincterotomy (incidence 11.1%; 95% CI, 9.2%–13.4%), inadvertent cannulation of the pancreatic duct, including with a guidewire (incidence 12.8%; 95% CI, 9.3%–17.3%), and any degree of pancreatic contrast injection (incidence 11.7%; 95% CI, 9.2%–14.6%) each resulted in high rates of PEP.

An indication of choledocholithiasis resulted in higher rates of clinically significant bleeding compared with an indication of malignant obstruction (2.6%; 95% CI, 1.7%–4.0% vs 0.7%; 95% CI, 0.5%–1.0%). This was the opposite for cholangitis, in which patients with malignant obstruction had an incidence of 8.7% (95% CI, 6.0%–12.6%) vs 2.3% (95% CI, 1.5%–3.5%) for patients with suspected choledocholithiasis; however, the malignant obstruction group's incidence may have been confounded by the placement of biliary stents, a group with a 9.8% incidence of cholangitis (95% CI, 7.0%–13.4%). There were insufficient study data to inform subgroup analyses for other AEs.

Sensitivity Analyses and Meta-regression

Results of sensitivity analyses are provided in Table 2. No significant differences in AE rates were demonstrated according to study location, mean enrollment dates, multi- or single-center status, or definitions used. Univariable meta-regression demonstrated that differences in mean age of participants may have contributed to heterogeneity in estimates of PEP and cholangitis. Despite adjustments for proportion of male patients, trainee involvement, and mean

Table 1. Pooled Incidences of Major ERCP-related AEs in All-Comers and First-Time Patients From RCTs and Large Observational Studies

	RCTs, all-comers	Observational studies, all-comers	Comparison (*P* value)	Overall incidence, all-comers	RCTs, first-time patients	Observational studies, first-time patients	Comparison (*P* value)	Overall incidence, first-time patients
PEP								
Incidence % (95% CI)	5.5 (4.5–6.7	4.1 (3.5–4.7)	<.01	**4.6 (4.0–5.1)**	7.0 (6.3–7.7)	5.2 (4.2–6.4)	<.01	**6.5 (5.9–7.1)**
Total patients (no. studies)	13,566 (66)	279,812 (77)		**293,378 (143)**	39,984 (124)	48,825 (31)		**88,809 (155)**
Heterogeneity % (I^2)	79	98		**96**	80	94		**89**
Clinically significant bleeding								
Incidence % (95% CI)	1.6 (1.1–2.1)	1.4 (1.1–1.8)	.69	**1.5 (1.2–1.7)**	1.9 (1.6–2.3)	1.9 (1.2–3.2)	.70	**1.9 (1.6–2.3)**
Total patients (no. studies)	8228 (38)	221,427 (52)		**229,655 (90)**	31,903 (101)	23,014 (19)		**54,917 (120)**
Heterogeneity % (I^2)	64	96		**93**	71	97		**88**
Cholangitis								
Incidence % (95% CI)	4.5 (3.0–6.6)	1.5 (1.1–2.2)	<.01	**2.5 (1.9–3.3)**	2.0 (1.6–2.5)	1.7 (1.1–2.4)	.27	**1.9 (1.6–2.4)**
Total patients (no. studies)	5713 (37)	115,906 (34)		**121,619 (71)**	22,494 (68)	15,343 (13)		**37,837 (81)**
Heterogeneity % (I^2)	79	97		**96**	77	83		**78**
Cholecystitis								
Incidence % (95% CI)	1.4 (1.1–1.9)	0.5 (0.2–1.3)	<.01	**0.8 (0.5–1.2)**	1.0 (0.6–1.9)	1.5 (0.7–2.8)	.73	**1.1 (0.6–1.9)**
Total patients (no. studies)	612 (7)	7187 (5)		**7799 (12)**	5895 (17)	607 (1)		**6,502 (18)**
Heterogeneity % (I^2)	0	60		**39**	69	N/A		**67**
Perforation								
Incidence % (95% CI)	0.8 (0.5–1.2)	0.5 (0.4–0.6)	.02	**0.5 (0.4–0.6)**	0.7 (0.6–0.8)	0.4 (0.3–0.6)	<.01	**0.6 (0.5–0.7)**
Total patients (no. studies)	6700 (28)	299,678 (54)		**306,378 (82)**	25,582 (84)	16,656 (14)		**42,238 (98)**
Heterogeneity % (I^2)	18	93		**90**	17	40		**30**
Death attributable to ERCP								
Incidence % (95% CI)	0.5 (0.3–0.8)	0.1 (0.1–0.2)	<.01	**0.2 (0.1–0.3)**	0.3 (0.3–0.4)	0.1 (0.1–0.2)	<.01	**0.3 (0.2–0.4)**
Total patients (no. studies)	5245 (17)	42,283 (21)		**47,528 (38)**	12,783 (38)	6,770 (7)		**19,553 (45)**
Heterogeneity % (I^2)	0	46		**44**	0	0		**0**

NOTE. Boldface values represent the overall incidences (combined between RCT and observational studies).
N/A, not applicable.

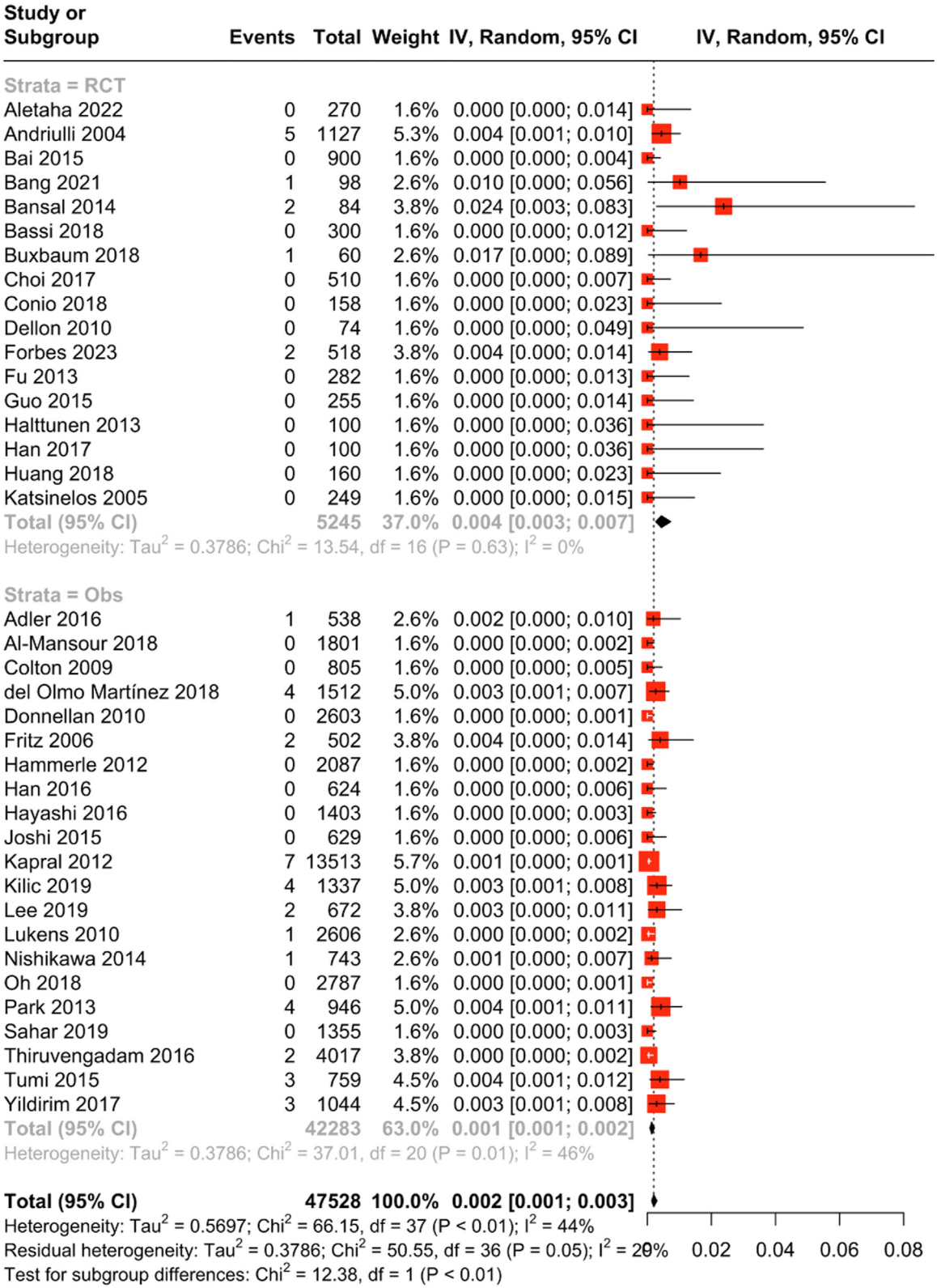


Figure 2. Forest plots demonstrating the risk of death attributable to ERCP in all-comers (data from randomized trials and observational studies shown separately, then pooled).

age of participants, substantial heterogeneity remained in incidences of AEs (Supplementary Table 8a). Meta-regression demonstrated no significant changes in the incidence of PEP over time based on mean RCT study recruitment dates between 2000 and 2023 for either the overall cohort (average annual percent change 0.06; 95%

Table 2. Pooled Risks of Major ERCP-related AEs Among Clinically Relevant Subgroups and Across Relevant Methodological Considerations and RRs of AEs for Clinically Relevant Pairwise Comparisons Using Pooled Data

Subgroup and sensitivity analyses				
Subgroup	Incidence from RCTs % (95% CI) No. of patients Heterogeneity % (I^2)	Incidence from observational studies % (95% CI) No. of patients Heterogeneity % (I^2)	Comparison (P value)	Overall incidence (pooled across study designs) % (95% CI) No. of patients Heterogeneity % (I^2)
Pancreatitis				
Demographics and procedural indications				
Female sex	10.0 (7.8–12.6) 4899 80	6.4 (5.4–7.6) 35,045 91	<.01	7.5 (6.4–8.6) 39,944 91
Male sex	6.3 (4.8–8.2) 4958 77	4.8 (4.0–5.6) 34,933 86	.07	5.2 (4.5–6.0) 39,891 85
Choledocholithiasis	6.3 (4.9–8.0) 4286 61	5.3 (3.7–7.6) 14,473 93	.44	5.8 (4.7–7.1) 18,759 86
Malignant obstruction	5.0 (3.6–6.9) 2776 62	5.2 (3.2–8.2) 8278 91	.90	5.1 (3.9–6.5) 11,054 82
Benign stricture	5.2 (1.2–19.8) 568 74	8.4 (3.9–17.3) 2426 90	.38	7.4 (4.1–12.8) 2994 88
Suspected or confirmed sphincter of Oddi dysfunction	15.6 (12.4–19.5) 1,719 48	16.9 (10.5–26.0) 305 47	.72	15.9 (12.1–19.1) 2,024 46
Pancreatic indications/ interventions	11.9 (3.8–31.8) 153 49	14.5 (1.2–71.2) 106 69	.75	12.8 (6.3–24.2) 259 52
Procedural parameters				
Sphincterotomy (with CBD guidewire) performed	5.2 (3.3–8.0) 1293 51	N/A	N/A	5.2 (3.3–8.0) 1293 51
Sphincteroplasty performed (of intact sphincter)	6.9 (4.5–10.5) 687 27	N/A	N/A	6.9 (4.5–10.5) 687 27
Sphincterotomy and sphincteroplasty performed	4.5 (2.8–7.2) 564 0	N/A	N/A	4.5 (2.8–7.2) 564 0
High-risk ERCP[a]	11.0 (9.4–12.9) 14,737 80	N/A	N/A	11.0 (9.4–12.9) 14,737 80
Difficult cannulation[b]	11.4 (9.2–13.9) 5313 69	N/A	N/A	11.4 (9.2–13.9) 5313 69
No pre-cut sphincterotomy performed on native papilla	7.2 (5.9-8.9) 7436 81	4.4 (3.2-5.8) 31,857 94	<.01	5.5 (4.5-6.7) 39,293 92
Pre-cut sphincterotomy performed on native papilla	14.1 (10.5-18.7) 895 40	9.8 (7.7-12.3) 3826 77	.04	11.1 (9.2-13.4) 4721 71
Pancreatic duct not cannulated	5.5 (3.7-7.9) 1328 32	4.4 (3.1-6.2) 5886 80	.28	4.8 (3.9-5.9) 7214 70
Pancreatic duct cannulated (with guidewire)	14.2 (6.4–28.4) 679 77	12.0 (8.4–16.9) 2193 83	.61	12.8 (9.3-17.3) 2872 84

Table 2. Continued

Subgroup and sensitivity analyses				
Subgroup	Incidence from RCTs % (95% CI) No. of patients Heterogeneity % (I^2)	Incidence from observational studies % (95% CI) No. of patients Heterogeneity % (I^2)	Comparison (P value)	Overall incidence (pooled across study designs) % (95% CI) No. of patients Heterogeneity % (I^2)
Pancreatic duct not injected	7.0 (4.9-10.1) 3061 85	4.5 (3.8-5.2) 12,953 68	.01	5.3 (4.4-6.5) 16,014 86
Pancreatic duct opacified (to any extent) with contrast	15.8 (11.5-21.2) 818 48	9.6 (7.1-13.0) 4431 81	.01	11.7 (9.2-14.6) 5249 80
Methodologic considerations				
Primarily North or South American setting(s)	5.6 (4.4–7.1) 7408 86	N/A	N/A	5.6 (4.4–7.1) 7408 86
Primarily European setting(s)	5.9 (4.5–7.6) 9523 80	N/A	N/A	5.9 (4.5–7.6) 9523 80
Primarily Asian-Pacific setting(s)	7.1 (6.3–8.0) 35,215 82	N/A	N/A	7.1 (6.3–8.0) 35,215 82
Median recruitment date between 2000 and 2007	6.3 (5.3–7.5) 13,622 70	N/A	N/A	6.3 (5.3–7.5) 13,622 70
Median recruitment date between 2008 and 2014	6.7 (5.7–7.8) 18,672 81	N/A	N/A	6.7 (5.7–7.8) 18,672 81
Median recruitment date between 2015 and 2023	6.2 (5.2–7.5) 16,850 81	N/A	N/A	6.2 (5.2–7.5) 16,850 81
Multicenter studies	5.9 (5.0–6.8) 29,101 85	N/A	N/A	5.9 (5.0–6.8) 29,101 85
Single-center studies	7.2 (6.3–8.1) 25,428 79	N/A	N/A	7.2 (6.3–8.1) 25,428 79
Consensus/ Lexicon definition employed for AEs[25,26]	6.5 (5.9–7.1) 39,332 76	N/A	N/A	6.5 (5.9–7.1) 39,332 76
Non-consensus, unclear, or absent definition of AEs	6.5 (5.2–8.0) 14,218 84	N/A	N/A	6.5 (5.2–8.0) 14,218 84
Bleeding				
Demographics and procedural indications				
Choledocholithiasis	2.6 (1.7–4.0) 2742 77	N/A	N/A	2.6 (1.7–4.0) 2742 77
Malignant obstruction	0.7 (0.5–1.0) 1155 0	N/A	N/A	0.7 (0.5–1.0) 1155 0
Procedural parameters				
Sphincterotomy (with CBD guidewire) performed	3.8 (2.3–6.2) 751 0	N/A	N/A	3.8 (2.3–6.2) 751 0
Sphincteroplasty performed (of intact sphincter)	0.9 (0.5–1.5) 836 0	N/A	N/A	0.9 (0.5–1.5) 836 0

Table 2. Continued

Subgroup and sensitivity analyses				
Subgroup	Incidence from RCTs % (95% CI) No. of patients Heterogeneity % (I^2)	Incidence from observational studies % (95% CI) No. of patients Heterogeneity % (I^2)	Comparison (P value)	Overall incidence (pooled across study designs) % (95% CI) No. of patients Heterogeneity % (I^2)
Sphincterotomy and sphincteroplasty performed	2.4 (1.6–3.4) 534 0	N/A	N/A	2.4 (1.6–3.4) 534 0
Pre-cut sphincterotomy performed	2.9 (0.8–9.7) 238 0	N/A	N/A	2.9 (0.8–9.7) 238 0
Methodologic considerations				
Primarily North or South American setting(s)	2.0 (1.6–2.6) 7886 25	N/A	N/A	2.0 (1.6–2.6) 7886 25
Primarily European setting(s)	2.6 (1.9–3.6) 8814 83	N/A	N/A	2.6 (1.9–3.6) 8814 83
Primarily Asian-Pacific setting(s)	1.5 (1.2–1.9) 22,865 73	N/A	N/A	1.5 (1.2–1.9) 22,865 73
Median recruitment date between 2000 and 2007	2.0 (1.5–2.8) 7064 47	N/A	N/A	2.0 (1.5–2.8) 7064 47
Median recruitment date between 2008 and 2014	1.8 (1.4–2.4) 16,302 84	N/A	N/A	1.8 (1.4–2.4) 16,302 84
Median recruitment date between 2015 and 2023	1.8 (1.5–2.3) 15,651 47	N/A	N/A	1.8 (1.5–2.3) 15,651 47
Multicenter studies	1.5 (1.1–1.8) 26,751 77	N/A	N/A	1.5 (1.1–1.8) 26,751 77
Single-center studies	2.4 (2.0–2.9) 13,472 68	N/A	N/A	2.4 (2.0–2.9) 13,472 68
Consensus/ Lexicon definition used for AEs[25,26]	1.7 (1.4–2.2) 19,705 76	N/A	N/A	1.7 (1.4–2.2) 19,705 76
Non-consensus, unclear, or absent definition of AEs	2.0 (1.6–2.5) 20,621 74	N/A	N/A	2.0 (1.6–2.5) 20,621 74
Cholangitis				
Procedural indications				
Choledocholithiasis	2.3 (1.5–3.5) 2236 49	N/A	N/A	2.3 (1.5–3.5) 2236 49
Malignant obstruction	8.7 (6.0–12.6) 1539 77	N/A	N/A	8.7 (6.0–12.6) 1539 77
Procedural parameters				
Biliary stent(s) placed	9.8 (7.0–13.4) 2155 76	N/A	N/A	9.8 (7.0–13.4) 2155 76
Methodologic considerations				
Primarily North or South American setting(s)	1.8 (1.0–3.3) 3872 79	N/A	N/A	1.8 (1.0–3.3) 3872 79

Table 2. Continued

Subgroup and sensitivity analyses				
Subgroup	Incidence from RCTs % (95% CI) No. of patients Heterogeneity % (I^2)	Incidence from observational studies % (95% CI) No. of patients Heterogeneity % (I^2)	Comparison (P value)	Overall incidence (pooled across study designs) % (95% CI) No. of patients Heterogeneity % (I^2)
Primarily European setting(s)	2.7 (1.5–4.9) 6656 88	N/A	N/A	2.7 (1.5–4.9) 6656 88
Primarily Asian-Pacific setting(s)	2.7 (2.1–3.5) 17,685 78	N/A	N/A	2.7 (2.1–3.5) 17,685 78
Median recruitment date between 2000 and 2007	2.4 (1.5–3.8) 5294 74	N/A	N/A	2.4 (1.5–3.8) 5294 74
Median recruitment date between 2008 and 2014	2.3 (1.6–3.4) 11,828 83	N/A	N/A	2.3 (1.6–3.4) 11,828 83
Median recruitment date between 2015 and 2023	2.8 (1.9–4.1) 9485 83	N/A	N/A	2.8 (1.9–4.1) 9485 83
Multicenter studies	2.7 (2.0–3.7) 19,055 86	N/A	N/A	2.7 (2.0–3.7) 19,055 86
Single-center studies	2.4 (1.7–3.3) 9158 73	N/A	N/A	2.4 (1.7–3.3) 9158 73
Consensus/ Lexicon definition used for AEs[25,26]	2.2 (1.7–2.8) 11,860 70	N/A	N/A	2.2 (1.7–2.8) 11,860 70
Non-consensus, unclear, or absent definition of AEs	3.1 (2.2–4.2) 16,972 87	N/A	N/A	3.1 (2.2–4.2) 16,972 87

Pairwise analyses				
Comparison	RR of event (95% CI)	Total number of patients	Number of studies	Heterogeneity % (I^2)
Pancreatitis				
Demographics and procedural indications				
Female sex (versus male sex)[c]	**1.39** (1.26–1.53)	39,944 / 39,891	47	36
Procedural parameters				
Double wire technique used (vs continued standard attempts)	1.75 (0.63–4.86)	317 / 1158	3	49
Double wire technique used (vs pre-cut sphincterotomy)	1.26 (0.79–2.02)	156 / 159	3	0
Pre-cut sphincterotomy performed (vs not)[c]	**2.07** (1.71–2.50)	4010 / 40,053	32	61
Pancreatic duct inadvertently cannulated (vs not)[c]	**3.26** (2.35–4.50)	2864 / 8217	15	70
Pancreatic duct opacified to any extent (vs not)[c]	**2.27** (1.93–2.68)	5185 / 19,302	28	47
Sphincterotomy performed (vs sphincteroplasty)	0.85 (0.38–1.93)	1682 / 653	9	54
Sphincterotomy alone performed (vs sphincterotomy with sphincteroplasty)	1.03 (0.76–1.38)	614 / 564	6	0

Table 2. Continued

Comparison	RR of event (95% CI)	Total number of patients	Number of studies	Heterogeneity % (I^2)
Pairwise analyses				
Interventions to mitigate the risk of PEP				
Rectal NSAIDs given (vs placebo or none)	**0.49** (0.38–0.63)	3471 / 3336	22	49
Prophylactic pancreatic stent placed (vs none)	**0.56** (0.43–0.72)	1845 / 1846	12	25
Aggressive prolonged intravenous fluids given (vs standard fluids or none)	**0.50** (0.33–0.75)	1412 / 1301	9	39
Bleeding				
Procedural parameters				
Pre-cut sphincterotomy performed (vs not)[c]	2.17 (0.04–107.16)	238 / 235	3	1
Sphincterotomy performed (vs sphincteroplasty)[c]	3.03 (0.22–41.97)	332 / 326	3	0
Sphincterotomy alone performed (vs sphincterotomy with sphincteroplasty)[c]	1.98 (0.96–4.08)	468 / 490	5	0

NOTE. Boldface values in the pairwise analyses indicate statistically significant findings.
CBD, common bile duct; N/A, not applicable; NSAID, nonsteroidal anti-inflammatory drug.
[a]High-risk ERCP represents a preselected cohort of patients at higher risk of post-ERCP pancreatitis according to study-specific eligibility criteria, based on patient- and/or procedure-related predictors. These are provided in the Supplementary Materials.
[b]Study-specific definitions of difficult cannulation provided in the Supplementary Materials.
[c]Pairwise comparisons of pooled observational data and "observational-type data from RCTs" given that patients were not randomized according to this variable.

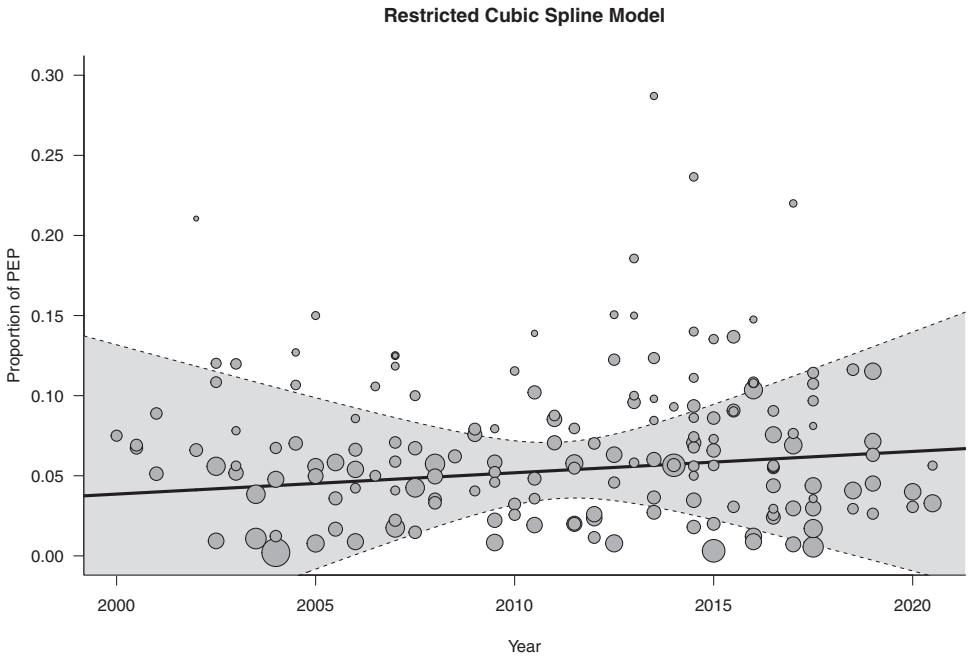


Figure 3. Meta-regression analysis of PEP incidence over time according to median recruitment dates for randomized trials (linear model).

CI, −0.27 to 0.39, Figure 3), all-comers only, or first-time patients only, and there was no evidence of a nonlinear relationship from cubic spline analyses (Supplementary Table 8b).

Pairwise Analyses

Pairwise analyses from RCTs are summarized in Table 2, with forest plots provided in Supplementary Figures 86 to 100. Female patients (RR, 1.39; 95% CI, 1.26–1.53 compared with male patients) had a higher risk of PEP. In terms of intraprocedural techniques, patients undergoing pre-cut sphincterotomy (RR, 2.07; 95% CI, 1.71–2.50 compared with none), inadvertent pancreatic duct cannulation (RR, 3.26; 95% CI, 2.35–4.50 compared with none), and pancreatic contrast injection (RR, 2.27; 95% CI, 1.93–2.68 vs none) all had higher risks of PEP (Figure 4). Rectal nonsteroidal anti-inflammatory drugs (RR, 0.49; 95% CI, 0.38–0.63 vs placebo), placement of a prophylactic pancreatic stent (RR, 0.56; 95% CI, 0.43–0.72 compared with none), and administration of aggressive prolonged intravenous hydration (RR, 0.50; 95% CI, 0.33–0.75 vs standard fluids or none) were all efficacious in mitigating PEP (Figure 5). There were no significant comparisons for bleeding and insufficient data to inform pairwise analyses for any other AEs.

Risk of Bias

Detailed risk of bias assessments for RCTs and observational studies are provided in Supplementary Tables 9 and 10, respectively. Overall, 67.8% of RCTs were deemed to be at "low risk" of bias, 20.9% had "some concerns" regarding bias, and 11.3% were deemed to be at "high risk," whereas 61.8% of observational studies were at low risk of bias (8 or 9 on the Newcastle-Ottawa Scale), whereas 38.2% were at moderate or high risk (7 or lower). Specifically, there were concerns with adequacy of follow-up in observational studies. There was visual and statistical evidence of publication bias for most main outcomes, with results of Egger's tests and funnel plots provided in Supplementary Table 11 and Supplementary Figures 101 to 124, respectively.

Certainty of Evidence

A GRADE summary of findings table summarizing the certainty of evidence for pairwise comparisons from RCTs is provided in Supplementary Table 12. Certainty of evidence ranged from very low to high, with the main issues resulting in down-grading being substantial statistical heterogeneity, imprecision, and use of "observational-type data from RCTs" as discussed.

Discussion

In this meta-analysis, we included data from 380 original studies comprising more than 2 million unique patients. There are several key findings. The rate of PEP in all-comers was 5.5% from more than 60 RCTs, and the rate of PEP in first-time ERCP patients was 7.0%, with both estimates being significantly higher than their corresponding values from observational studies (4.1% and 5.2%, respectively). The rates of bleeding, cholangitis, cholecystitis, and perforation were also identified from both RCTs and observational studies. Several risk factors for PEP and bleeding were identified, and several interventions demonstrated efficacy in mitigating PEP from our up-to-date analyses, including rectal nonsteroidal anti-inflammatory drugs, prophylactic pancreatic stent placement, and aggressive prolonged intravenous fluids.

ERCP is complex and is associated with the highest AE risks of all commonly performed gastrointestinal endoscopies. Describing the incidences of and elucidating the risk factors for these AEs has been the primary focus of thousands of observational studies in our field, whereas determining efficacious interventions to mitigate their occurrence has been the focus of hundreds more RCTs. Since the initial descriptions of endoscopic papillary cannulation and development of side-viewing duodenoscopes in the 1960s[35,36] and subsequent reports of the first pancreaticobiliary therapeutics in the 1970s,[37,38] ERCP has evolved considerably. Of note, despite more than 2 full generations passing since its inception, the AE profile of ERCP remains high. In a seminal study recruiting more than 2300 patients with native papillae from the early 1990s, Freeman et al[39] demonstrated a 5.4% incidence of PEP and a 2.0% incidence of clinically significant bleeding. From large observational studies performed since 2000, we determined a PEP incidence of 5.2% in more than 48,000 patients with native papillae and a bleeding incidence of 1.9%—rates that are remarkably similar. To contextualize this lack of change using examples across similar timeframes from surgical and endoscopic literature, the incidence of bile duct injury with laparoscopic cholecystectomy has decreased 3-fold between 1994 to 2014,[40] and clinically significant bleeding following routine colonic polypectomy has decreased from 2% to 3% in the late 1990s to now being such a rare outcome that it is difficult to measure and report.[41,42] Reasons for these improvements are due to advances in available technologies and techniques responsible for both, guided by increasing user experience and high-quality research. Why is it, then, that incidences of post-ERCP AEs have remained static over similar timeframes despite corresponding improvements in endoscopy technology and techniques? The answer to this question is likely nuanced and multifactorial.

Given the very nature of the procedure and its anatomical target(s), it could be a reality that no matter how much our understanding of ERCP AE mechanisms and risk factors improves, a significant risk of AEs will remain. Even in cases of routine biliary indications for ERCP, the proximity of the pancreatic duct opening to the biliary orifice means that patients will always be at risk of PEP—this is evidenced by cases where, even despite a simple cannulation in a patient with no traditional risk factors, PEP ensues. There is much we have yet to understand regarding the mechanisms of PEP, and much could be explained by anatomical variations and/or abnormalities beyond our control. Similarly, explaining persistently high rates of clinically significant bleeding is multifactorial and could relate in part to

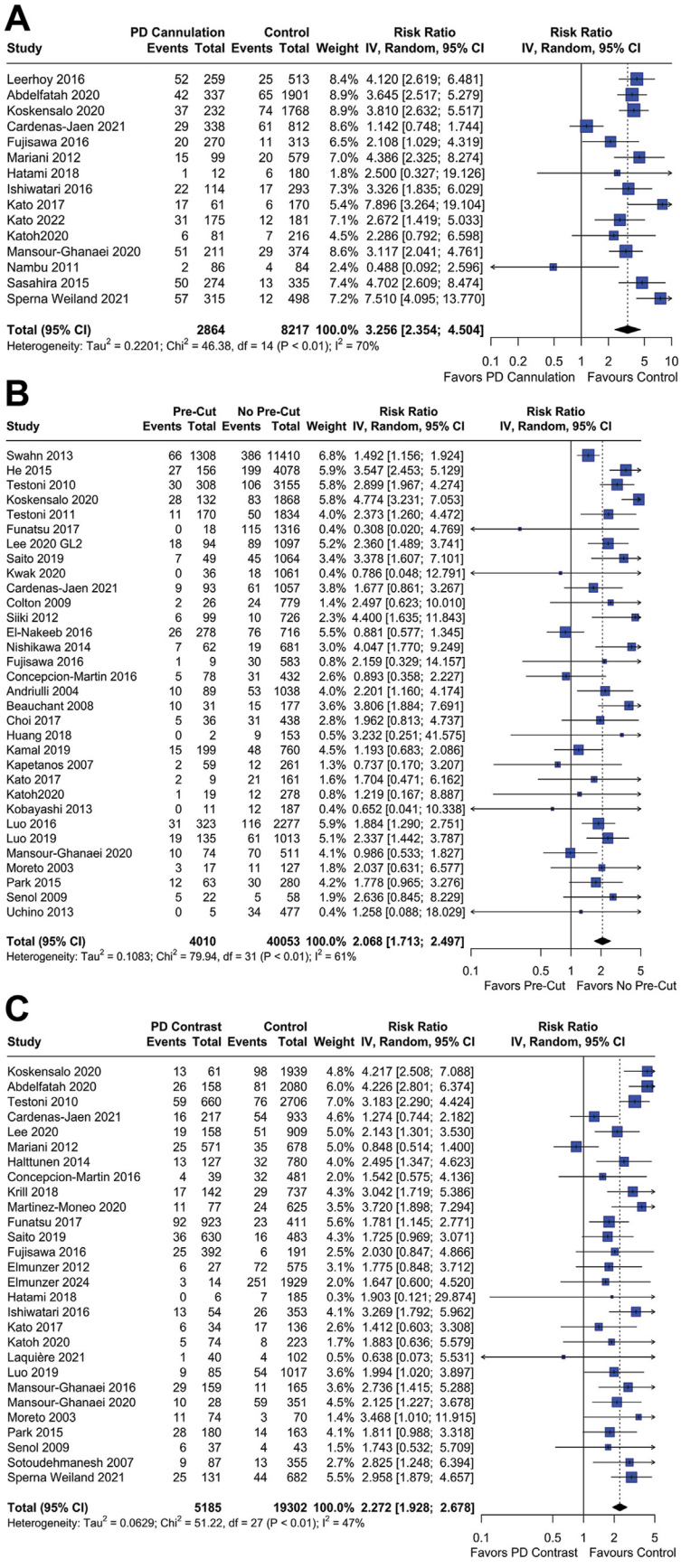
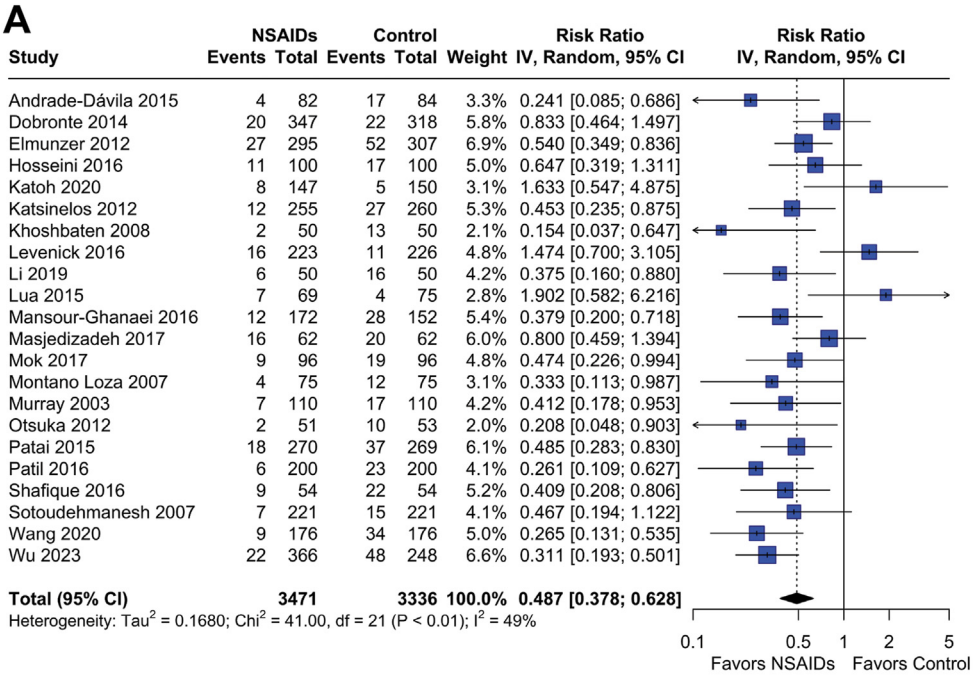
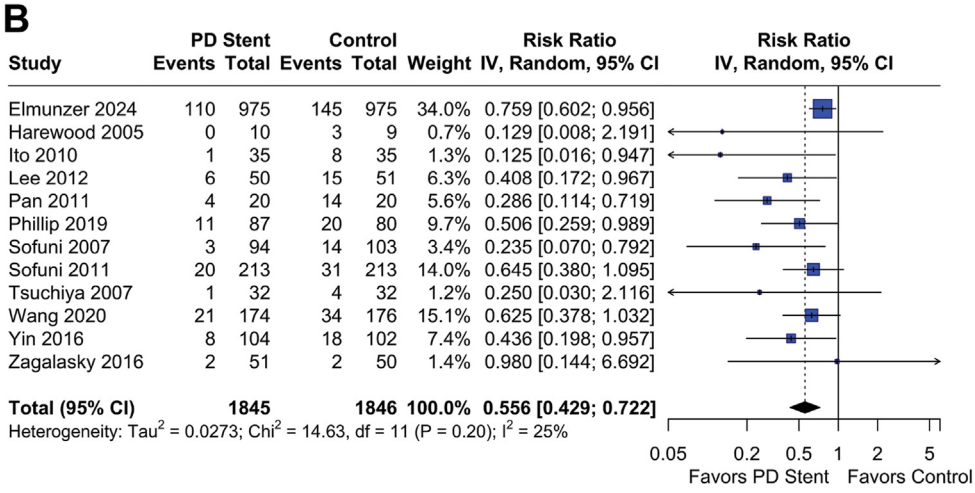


Figure 4. Forest plots demonstrating the relative risks of PEP with commonly performed intra-procedural techniques: (*A*) pancreatic duct cannulation (vs none); (*B*) pre-cut sphincterotomy (vs none); and (*C*) pancreatic duct opacification (any vs none).





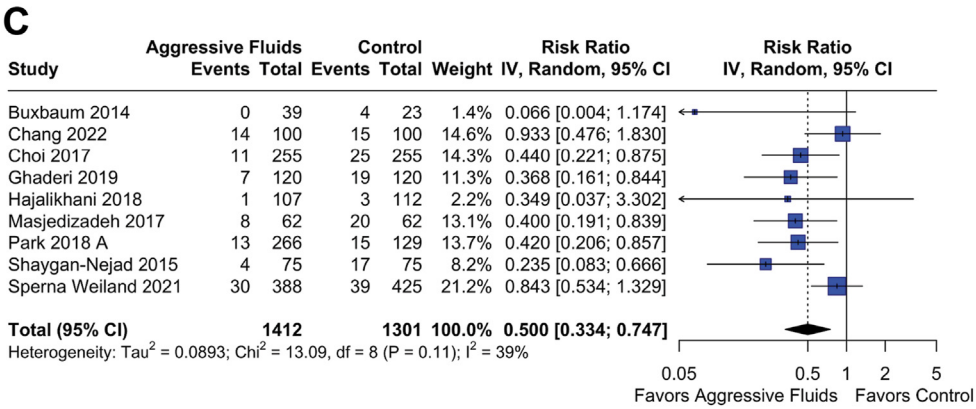


Figure 5. Forest plots demonstrating the RR reductions in PEP with interventions studied to mitigate its effect in randomized trials: (*A*) nonsteroidal anti-inflammatory drugs given per rectum (vs none or placebo); (*B*) pancreatic duct stenting (vs none); and (*C*) aggressive prolonged intravenous fluids (vs none or standard fluids).

suboptimal peri-procedural management of antithrombotic agents, especially antiplatelet medications.[43] These factors in addition to the known challenging learning curve associated with ERCP training[1–3] can partially explain persistently high AE rates. Second, ERCP today is performed almost exclusively for therapeutic indications, whereas a generation ago, a typical ERCP case mix would comprise both diagnostic and therapeutic indications.[7] Thus, it is

plausible that background improvements in technology, available equipment, and techniques could be offset by higher baseline risks of interventions we are now performing in far greater proportions. Third, it is possible that current ranges of ERCP AEs encountered are potentially the result of between-provider variability, with both low endoscopist and low center volumes being directly correlated.[10,11] Although this variability is disconcerting, these data suggest that there are potential interventions that could modify outcomes, and these could include efforts to re-centralize provision of ERCP and/or distribution of audit and feedback measures (ie, report cards) or educational interventions, that, although proven beneficial in other endoscopic domains,[44,45] remain relatively understudied in ERCP.[46] Researching such interventions is crucial given the substantial morbidity, mortality, and health care burden associated with ERCP AEs.

Being armed with precise estimates of ERCP-associated AE risk is crucial for any endoscopist performing the procedure for 3 main reasons. The first is to be able to engage in detailed informed consent discussions with patients on both global and specific risks. For this reason, we identified several important risk factors for specific AEs and reported subgroup-specific incidence estimates across these relevant patient- and procedure-related factors in addition to providing estimates of comparative risk. This will hopefully result in a clearer understanding of these higher-risk groups for both patients and endoscopists. The second reason this is important, beyond consent, is that endoscopists performing ERCP on such higher-risk patients should use evidence- and guideline-based strategies to mitigate their risk, wherever possible.[47–50] A final reason is that precise estimates in various contexts are also crucial for the estimation of sample sizes in the conduct of clinical research. Several systematic reviews and meta-analyses have been published regarding ERCP AEs; however, all have been more focused in scope, and none to our knowledge have attempted to synthesize available data on all common ERCP-related AEs. The largest reviews to date have focused on PEP alone. Our study is the most comprehensive review of ERCP-associated AEs performed to date.

We made the decision a priori to include data from both RCTs and large observational studies, and to report results both separately and pooled together, where appropriate. The rationale behind this decision is important. In our study, in which estimates of AE incidences were the primary outcome, one can argue that RCT data and observational data should be treated no differently, hence the decision to pool these data together in the reporting of our results. However, we also felt it was important to provide readers with statistical comparisons between incidence estimates derived from RCTs vs observational studies. It is well established that estimates derived from RCTs can vary considerably because of important differences in underlying populations included,[25] outcomes ascertainment approaches,[26] and event reporting,[27] among other factors. Conversely, observational data, especially when it comes to granular intraprocedural parameters in ERCP, can be skewed for numerous reasons, including self-reporting bias and retrospective bias, among others.[51] Indeed, statistically significant and clinically important differences existed between these groups in our results, likely pointing to considerable variability in underlying methodological approaches, as outlined previously. As might be expected, estimates of AE incidences were consistently higher from pooled RCT data. RCTs generally have more rigorous outcome definitions and protocolized follow-up, collectively increasing the likelihood of AE detection that is lacking in observational studies. However, trials also likely represent more complex case mixes than are seen in routine clinical practices owing to the higher likelihood of participation by academic (ie, tertiary or quaternary) centers. Thus, true incidences likely lie in between these 2 estimates, and it is crucial to have both to fully understand the spectrum of ERCP-associated AEs. Ultimately, these estimates are meant to provide a guide to endoscopists performing ERCP, who can then best estimate the risk for a given patient only after considering all patient-, procedure-, and endoscopist-related factors.

In addition to the preceding decision and associated rationale, our study has additional strengths. We made the decision to include only studies inclusive of ERCPs primarily performed after the year 2000. This is important given that (1) before 2000, ERCP volumes in the United States comprised large proportions of diagnostic procedures, whereas today, ERCP is performed almost exclusively for therapeutic indications,[7] and (2) procedures performed before 2000 are not representative of modern equipment and newer techniques—for example, it is of questionable relevance to cite AEs from a study performed more than 40 years ago to inform a discussion of risks with a patient in 2024. In addition, we performed meticulous data extraction and categorization of patients, procedural characteristics, and outcomes, allowing the identification of all-comers, patients undergoing initial ERCP (ie, those with native papillae), and high-risk patients with high degrees of confidence. Furthermore, we opted to include both arms of RCTs in our calculations of pooled AE incidences rather than placebo-only arms, as has been done in a prior meta-analysis.[12] We did this in an effort to have our results reflect accurate contemporary practices, both in terms of realistic variations in intraprocedural techniques performed and inconsistent usage of available interventions to mitigate PEP, which has been reliably demonstrated.[52,53] Finally, we analyzed and reported our results by relevant subgroups, which has not been done to this extent in prior evidence syntheses.[12,17] This is critical to help endoscopists offer tailored discussions of risks for patients undergoing ERCP and to assist them in making decisions regarding mitigating strategies and/or closer monitoring for patients at higher risk.

Our study also has limitations. First, there was substantial to considerable statistical heterogeneity in several of our pooled estimates. This is unsurprising given the large number of included studies, many of which were aimed at investigating the effects of interventions or at determining AEs within specific populations. We attempted to elucidate potential sources of this heterogeneity by performing several subgroup and sensitivity analyses. We also

performed univariable meta-regression, which demonstrated that differences in mean age of participants may have contributed to the heterogeneity observed in the estimates of pancreatitis and cholangitis. However, despite adjustments for sex and trainee involvement (in addition to mean age of participants), there still remained substantial heterogeneity in the incidence of AEs between included studies. Second, we elected to not include conference abstracts. We felt inclusion of gray literature would have introduced another potential source of heterogeneity because of absence of detailed methodology and outcome definitions. However, we acknowledge that although this decision may have reduced heterogeneity, it may have also introduced potential publication bias. Finally, although our decision to set a minimum threshold of patients for included observational studies was made to mitigate potential biases introduced due to small-study effects and our rare outcome, we acknowledge that we may not have captured smaller studies addressing some subgroups of interest, such as bleeding in those with liver disease or those on antithrombotic medications, AEs in patients with primary sclerosing cholangitis or in patients undergoing cholangioscopy, or AEs in patients undergoing ERCP with novel duodenoscope designs.

To conclude, our study provides contemporary data informing estimated incidences of all common ERCP-related AEs as well as estimates of the magnitude of risk associated with several clinically relevant patient- and procedure-related factors. These results are important to patients, endoscopists, and policy makers, to raise awareness and facilitate implementation of evidence-based interventions to mitigate risk.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2024.10.033.

References

1. Johnson G, Webster G, Boškoski I, et al. Curriculum for ERCP and endoscopic ultrasound training in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2021;53:1071–1087.
2. Wani S, Keswani RN, Han S, et al. Competence in endoscopic ultrasound and endoscopic retrograde cholangiopancreatography, from training through independent practice. Gastroenterology 2018;155:1483–1494.e7.
3. Wani S, Hall M, Wang AY, et al. Variation in learning curves and competence for ERCP among advanced endoscopy trainees by using cumulative sum analysis. Gastrointest Endosc 2016;83:711–719.e11.
4. Chandrasekhara V, Khashab MA, Muthusamy VR, et al. Adverse events associated with ERCP. Gastrointest Endosc 2017;85:32–47.
5. Dumonceau JM, Kapral C, Aabakken L, et al. ERCP-related adverse events: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy 2020;52:127–149.
6. Kroner PT, Bilal M, Samuel R, et al. Use of ERCP in the United States over the past decade. Endosc Int Open 2020;8:E761.
7. Ahmed M, Kanotra R, Savani GT, et al. Utilization trends in inpatient endoscopic retrograde cholangiopancreatography (ERCP): a cross-sectional US experience. Endosc Int Open 2017;5:E261–E271.
8. Forbes N, Coelho-Prabhu N, Al-Haddad MA, et al. Adverse events associated with EUS and EUS-guided procedures. Gastrointest Endosc 2022;95:16–26.e2.
9. Thiruvengadam NR, Saumoy M, Schaubel DE, et al. Rise in first-time ERCP for benign indications >1 year after cholecystectomy is associated with worse outcomes. Clin Gastroenterol Hepatol 2024;22:1618–1627.e4.
10. Teles de Campos S, Papaefthymiou A, Florou T, et al. Impact of center and endoscopist ERCP volume on ERCP outcomes: a systematic review and meta-analysis. Gastrointest Endosc 2023;98:306–315.e14.
11. Huang RJ, Barakat MT, Girotra M, et al. Unplanned hospital encounters after endoscopic retrograde cholangiopancreatography in 3 large North American states. Gastroenterology 2019;156:119–129.e3.
12. Akshintala VS, Kanthasamy K, Bhullar FA, et al. Incidence, severity, and mortality of post-ERCP pancreatitis: an updated systematic review and meta-analysis of 145 randomized controlled trials. Gastrointest Endosc 2023;98:1–6.e12.
13. Mutneja H, Vohra I, Go A, et al. Temporal trends and mortality of post-endoscopic retrograde cholangiopancreatography pancreatitis in the United States: a nationwide analysis. Endoscopy 2021;63:357–366.
14. Fagenholz PJ, Fernández-del Castillo C, Harris NS, et al. Direct medical costs of acute pancreatitis hospitalizations in the United States. Pancreas 2007;35:302–307.
15. Inamdar S, Sejpal DV, Ullah M, et al. Weekend vs. weekday admissions for cholangitis requiring an ERCP: comparison of outcomes in a national cohort. Am J Gastroenterol 2016;111:405–410.
16. Kochar B, Akshintala VS, Afghani E, et al. Incidence, severity, and mortality of post-ERCP pancreatitis: a systematic review by using randomized, controlled trials. Gastrointest Endosc 2015;81:143–149.e9.
17. Andriulli A, Loperfido S, Napolitano G, et al. Incidence rates of post-ERCP complications: a systematic survey of prospective studies. Am J Gastroenterol 2007;102:1781–1788.
18. Elmunzer BJ, Waljee AK, Elta GH, et al. A meta-analysis of rectal NSAIDs in the prevention of post-ERCP pancreatitis. Gut 2008;57:1262–1267.
19. Page MJ, McKenzie JE, Kirkham J, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Database Syst Rev 2014;10:MR000035.
20. Golder S, Loke YK, Bland M. Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. PLoS Med 2011;8:e1001026.

21. Rivas A, Pherwani S, Mohamed R, et al. ERCP-related adverse events: incidence, mechanisms, risk factors, prevention, and management. Expert Rev Gastroenterol Hepatol 2023;17:1101–1116.
22. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.
23. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; 283:2008–2012.
24. Forbes N, Leontiadis G, Vaska M, et al. Adverse events associated with endoscopic retrograde cholangiopancreatography: protocol for a systematic review and meta-analysis. BMJ Open 2021;11:e053302.
25. He J, Morales DR, Guthrie B. Exclusion rates in randomized controlled trials of treatments for physical conditions: a systematic review. Trials 2020;21:228.
26. Mc Cord KA, Ewald H, Agarwal A, et al. Treatment effects in randomised trials using routinely collected data for outcome assessment versus traditional trials: meta-research study. BMJ 2021;372:n450.
27. Westergren T, Narum S, Klemp M. Biases in reporting of adverse effects in clinical trials, and potential impact on safety assessments in systematic reviews and therapy guidelines. Basic Clin Pharmacol Toxicol 2022; 131:465–473.
28. Cotton PB, Lehman G, Vennes J, et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. Gastrointest Endosc 1991; 37:383–393.
29. Cotton PB, Eisen GM, Aabakken L, et al. A lexicon for endoscopic adverse events: report of an ASGE workshop. Gastrointest Endosc 2010;71:446–454.
30. Forbes N, Elmunzer BJ, Keswani RN, et al. Consensus-based development of a causal attribution system for post-ERCP adverse events. Gut 2022;71:1963–1966.
31. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.
32. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses, 2018. Available at: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed March 21, 2024.
33. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–634.
34. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–926.
35. McCune WS, Shorb PE, Moscovitz H. Endoscopic cannulation of the ampulla of vater: a preliminary report. Ann Surg 1968;167:752–756.
36. Oi I, Takemoto T, Kondo T. Fiberduodenoscope: direct observation of the papilla of Vater. Endoscopy 1969; 3:101–103.
37. Zimmon DS, Falkenstein DB, Kessler RE. Endoscopic papillotomy for choledocholithiasis. N Engl J Med 1975; 293:1181–1182.
38. Laurence BH, Cotton PB. Decompression of malignant biliary obstruction by duodenoscopic intubation of bile duct. Br Med J 1980;280:522–523.
39. Freeman ML, Nelson DB, Sherman S, et al. Complications of endoscopic biliary sphincterotomy. N Engl J Med 1996;335:909–918.
40. Pucher PH, Brunt LM, Davies N, et al. Outcome trends and safety measures after 30 years of laparoscopic cholecystectomy: a systematic review and pooled data analysis. Surg Endosc 2018;32:2175–2183.
41. Parra-Blanco A, Kaminaga N, Kojima T, et al. Colonoscopic polypectomy with cutting current: is it safe? Gastrointest Endosc 2000;51:676–681.
42. Shinozaki S, Kobayashi Y, Hayashi Y, et al. Efficacy and safety of cold versus hot snare polypectomy for resecting small colorectal polyps: systematic review and meta-analysis. Dig Endosc 2018;30:592–599.
43. Bishay K, Ruan Y, Barkun AN, et al. Incidence, predictors, and outcomes of clinically significant post-endoscopic retrograde cholangiopancreatography bleeding: a contemporary multicenter study. Am J Gastroenterol 2024;119:2317–2325.
44. Bishay K, Causada-Calo N, Scaffidi MA, et al. Associations between endoscopist feedback and improvements in colonoscopy quality indicators: a systematic review and meta-analysis. Gastrointest Endosc 2020; 92:1030–1040.e9.
45. Causada-Calo NS, Gonzalez-Moreno EI, Bishay K, et al. Educational interventions are associated with improvements in colonoscopy quality indicators: a systematic review and meta-analysis. Endosc Int Open 2020; 8:E1321–E1331.
46. Coté GA, Elmunzer BJ, Forster E, et al. Development of an automated ERCP Quality Report Card using structured data fields. Tech Innov Gastrointest Endosc 2021; 23:129–138.
47. Buxbaum JL, Freeman M, Amateau SK, et al. American Society for Gastrointestinal Endoscopy guideline on post-ERCP pancreatitis prevention strategies: summary and recommendations. Gastrointest Endosc 2023;97:153–162.
48. Buxbaum JL, Freeman M, Amateau SK, et al. American Society for Gastrointestinal Endoscopy guideline on post-ERCP pancreatitis prevention strategies: methodology and review of evidence. Gastrointest Endosc 2023; 97:163–183.e40.
49. Elmunzer BJ, Foster LD, Serrano J, et al. Indomethacin with or without prophylactic pancreatic stent placement to prevent pancreatitis after ERCP: a randomised non-inferiority trial. Lancet 2024;403:450–458.
50. Elmunzer BJ, Scheiman JM, Lehman GA, et al. A randomized trial of rectal indomethacin to prevent post-ERCP pancreatitis. N Engl J Med 2012;366:1414.
51. Forbes N, Koury HF, Bass S, et al. Characteristics and outcomes of ERCP at a Canadian tertiary centre: initial results from a prospective high-fidelity biliary endoscopy registry. J Can Assoc Gastroenterol 2021;4:78–83.

52. Smith ZL, Elmunzer BJ, Cooper GS, et al. Real-world practice patterns in the era of rectal indomethacin for prophylaxis against post-ERCP pancreatitis in a high-risk cohort. Am J Gastroenterol 2020;115:934–940.
53. Avila P, Holmes I, Kouanda A, et al. Practice patterns of post-ERCP pancreatitis prophylaxis techniques in the United States: a survey of advanced endoscopists. Gastrointest Endosc 2020;91:568–573.e2.

Received August 8, 2024. Accepted October 8, 2024.

Correspondence
Address correspondence to: Nauzer Forbes, MD, MSc, FASGE, Division of Gastroenterology and Hepatology, University of Calgary, CWPH 6D19, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. e-mail: nauzer.forbes@ucalgary.ca.

CRediT Authorship Contributions
Kirles Bishay, MD, MSc (Investigation: Supporting; Supervision: Supporting; Writing – original draft: Lead; Writing – review & editing: Supporting)
Zhao Wu Meng, MD, MSc (Investigation: Supporting; Supervision: Supporting; Writing – original draft: Lead; Writing – review & editing: Supporting)
Rishad Khan, MD (Writing – original draft: Supporting; Writing – review & editing: Supporting)
Mehul Gupta, MD (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)
Yibing Ruan, PhD, MPH (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)
Marcus Vaska, MLIS (Investigation: Supporting; Writing – review & editing: Supporting)
Jordan Iannuzzi, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Dylan E. O'Sullivan, PhD (Investigation: Supporting; Writing – review & editing: Supporting)
Brittany Mah, BSc (Investigation: Supporting; Writing – review & editing: Supporting)
Amanda M. Henderson, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Arun C.R. Partridge, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Howard Guo, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Sunil Samnani, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Max DeMarco, BSc (Investigation: Supporting; Writing – review & editing: Supporting)
Yuhong Yuan, PhD (Methodology: Supporting; Writing – review & editing: Supporting)
B. Joseph Elmunzer, MD, MSc (Investigation: Supporting; Writing – review & editing: Supporting)
Rajesh N. Keswani, MD, MSc (Investigation: Supporting; Writing – review & editing: Supporting)
Sachin Wani, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Zachary L. Smith, DO (Investigation: Supporting; Writing – review & editing: Supporting)
Ronald J. Bridges, MD (Investigation: Supporting; Writing – review & editing: Supporting)
Steven J. Heitman, MD, MSc (Investigation: Supporting; Writing – review & editing: Supporting)
Robert J. Hilsden, MD, PhD (Investigation: Supporting; Writing – review & editing: Supporting)
Darren R. Brenner, PhD (Investigation: Supporting; Supervision: Supporting; Writing – review & editing: Supporting)
Grigorios I. Leontiadis, MD, PhD (Investigation: Supporting; Supervision: Supporting; Writing – review & editing: Supporting)
Nauzer Forbes, MD MSc (Conceptualization: Lead; Investigation: Lead; Supervision: Lead; Writing – original draft: Supporting; Writing – review & editing: Lead)

Conflicts of interest
The authors disclose no conflicts.

Data Availability
All source data in this study are available from the original publications. All analytic methods and study materials can be made available upon request on a case-by-case basis depending on the purpose. Such requests can go to the corresponding author.