



Published in final edited form as:

*Am J Gastroenterol.* 2013 November ; 108(11): . doi:10.1038/ajg.2013.217.

## Intraprocedural Quality in Endoscopic Retrograde Cholangiopancreatography: A Meta-Analysis

Anthony T. DeBenedet, MD<sup>1</sup>, B. Joseph Elmunzer, MD<sup>1</sup>, Sean T. McCarthy, MD<sup>1</sup>, Grace H. Elta, MD<sup>1</sup>, and Philip S. Schoenfeld, MD, MEd, MSc(Epi)<sup>1</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Department of Internal Medicine, University of Michigan Health System, Ann Arbor, Michigan, USA

### Abstract

**OBJECTIVES**—In 2006, the American College of Gastroenterology (ACG)/the American Society for Gastrointestinal Endoscopy (ASGE) Taskforce on Quality in Endoscopy published quality indicators for the major gastrointestinal procedures. Our primary aim was to use the published literature to assess current endoscopic retrograde cholangiopancreatography (ERCP) intraprocedural performance and compare it to the targets set by the ACG/ASGE taskforce. Our secondary aim was to determine whether performance varies across different health-care settings (academic and community), study designs (prospective and retrospective), and trainee participation.

**METHODS**—A PubMed and EMBASE literature search from 1/1/2006 to 2/1/2013 was conducted. Articles were selected based on title, abstract, full text, and reporting of success rates for the intraprocedural quality indicators. Success rates, represented as numerical proportions, were collected from each study. For each success rate, a standard error and a 95% confidence interval (CI) was calculated. A random-effects meta-analysis model was used to weight each study, and a cumulative, weighted success rate (or effect size) for each indicator was determined. Random-effects meta-regression was then used to examine the impact of study setting, design, and trainee involvement on each quality indicator.

**RESULTS**—A total of 8,005 articles were initially retrieved. Following the application of predefined criteria, 52 articles remained. The cumulative, weighted bile duct cannulation success rate was 89.3% (95% CI 0.866–0.919); pancreatic duct cannulation was 85.0% (95% CI 0.813–0.886); precut utilization rate was 10.5% (95% CI 0.087–0.123); common bile duct stone extraction rate was 88.3% (95% CI 0.825–0.941); and the rate of successful biliary stenting below the common bile duct bifurcation was 97.5% (95% CI 0.967–0.984). Subgroup analysis with meta-regression showed no statistically significant differences between academic and community settings, prospective and retrospective study designs, and trainee participation on success across

© 2013 by the American College of Gastroenterology

Correspondence: Anthony T. DeBenedet, MD, Division of Gastroenterology and Hepatology, Department of Internal Medicine, University of Michigan Health System, 3912 Taubman Center, 3912 Taubman Center, 1500 E. Medical Center Dr., SPC 5362, Ann Arbor, Michigan 48109-5362, USA. adebened@med.umich.edu.

**Specific author contributions** : Anthony T. DeBenedet—conception and design; analysis and interpretation of the data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; final approval of the article. B. Joseph Elmunzer—critical revision of the article for important intellectual content; final approval of the article. Sean T. McCarthy—analysis and interpretation of the data; critical revision of the article for important intellectual content; final approval of the article. Grace H. Elta—critical revision of the article for important intellectual content; final approval of the article. Philip S. Schoenfeld—conception and design; critical revision of the article for important intellectual content; final approval of the article.

**Potential competing interests**: None.

**Guarantor of the article**: Anthony T. DeBenedet, MD.

bile duct cannulation, precut utilization, and common bile duct stone extraction (insufficient observations/variance for pancreatic duct cannulation and biliary stent placement).

**CONCLUSIONS**—ERCP intraprocedural quality is in good standing. On the basis of this analysis, the two targets that could be potentially revised are precut utilization and biliary stenting. This analysis was confined to the published literature and therefore, in general, reflects the ERCP performance of institutions, primarily academic, that are conducting clinical research. Thus, it is difficult to generalize this performance assessment to the broader ERCP community as a whole.

## INTRODUCTION

Health-care quality is a complex, multifaceted concept. From a clinical viewpoint, this complexity arises because it is very difficult to know—with complete certainty—whether a quality indicator or list of indicators actually reflect the quality of the care being delivered; in other words, it is hard to prove the validity of any quality metric (1). Despite this fact, almost all health-care stakeholders—including the majority of physicians—believe that striving to better understand, measure, and assure health-care quality should be a top priority (2). This is especially relevant in light of the Affordable Care Act, where quality is a central tenet, as well as the Centers for Medicare & Medicaid Services' Physician Quality Reporting Initiative, which currently offers provider incentives for reporting quality data and in 2015 will implement payment adjustments for providers who do not satisfactorily report data on quality measures. The end goals of the quality movement in healthcare are to increase transparency, help providers achieve and maintain individual and collective excellence, and provide patients with optimal care.

In 2006, the American College of Gastroenterology (ACG) and the American Society for Gastrointestinal Endoscopy (ASGE) Taskforce on Quality in Endoscopy published quality metrics for the major gastrointestinal procedures: esophagogastroduodenoscopy; colonoscopy; endoscopic ultrasound; and endoscopic retrograde cholangiopancreatography (ERCP) (3,4). These metrics are intended to be measured by health systems and help guide endoscopic performance improvement initiatives.

This meta-analysis focuses on the quality metrics published for ERCP, specifically the intraprocedural quality indicators: (i) achievement of deep cannulation of the bile duct; (ii) achievement of deep cannulation of the main pancreatic duct; (iii) limiting the use of precut techniques for achieving ductal cannulation; (iv) successful extraction of common bile duct stones during first ERCP; and (v) successful biliary stent placement for biliary obstruction below the bifurcation of the common bile duct.

Our primary aim was to use the published literature to assess current ERCP intraprocedural performance and compare it to the targets set by the ACG/ASGE taskforce. Our secondary aim was to determine whether performance varies across different health-care settings (academic and community), study designs (prospective and retrospective), and trainee participation.

## METHODS

A PubMed and EMBASE literature search from 1/1/2006 to 2/1/2013 of studies published in English was conducted. This timeframe represents the period since the quality metrics were published, and it is most representative of current practice patterns. Search strings were constructed using terms that commonly describe the intraprocedural quality indicators. Articles were assessed in duplicate and independent fashion by two investigators (ATD, STM) and were selected based on title, abstract, full text, and reporting of success rates for the intra procedural quality indicators. Articles were not required to have a primary aim of

assessing ERCP intraprocedural quality. The references of all retrieved articles were scrutinized for additional studies that might have been missed by the search strategy.

Potentially relevant studies were assessed for the following inclusion criteria (as relevant for each intraprocedural quality indicator): patients > 18 years of age; consecutive patient selection; reporting of bile duct and main pancreatic duct cannulation success rate using standard cannulation techniques (catheters, papillotomes, guidewires in conjunction with catheters and papillotomes, and placement of pancreatic stent or guidewire to facilitate cannulation); reporting of precut utilization rate using standard precut techniques (needle knife papillotomy, suprapapillary puncture of the common bile duct (needle knife fistulotomy); reporting of common bile duct stone extraction success rate during the first ERCP session using standard techniques (biliary sphincterotomy, balloon dilation, basket extraction, and mechanical lithotripsy); and reporting of successful biliary stent placement for obstruction below the bifurcation. Studies were excluded if patients with altered anatomy were included in success rate calculations (i.e., prior pancreaticoduodenectomy, Billroth II anatomy, prior gastrojejunostomy, prior hepaticojejunostomy), or patients with duodenal obstruction were included in success rate calculations. Randomized controlled trials were also excluded, given their nonconsecutive patient selection process.

For each study included in our analysis, data were abstracted onto standardized electronic spreadsheets in duplicate and independent fashion by two investigators (ATD, STM). Numerical proportions representing the success rates for the intraprocedural quality indicators were collected from each study. Also, data on the geographical region where the study took place (country), the hospital setting for the procedures (academic, community, or both), the study design (prospective or retrospective), and trainee involvement were also collected. All discrepancies in data extraction were resolved by consensus after re-reviewing the study in question.

For each success rate, a standard error and a 95 % confidence interval (CI) was calculated. A random-effects meta-analysis model was used to weight each study, and a cumulative, weighted success rate (or effect size) for each indicator was determined. Random-effects meta-regression was then used to examine the impact of study setting, design, and trainee involvement on each quality indicator.

All analyses were conducted with Stata statistical software (Intercooled Stata, version 12; StataCorp LP, College Station, TX).

## RESULTS

A total of 8,005 articles were initially retrieved using relevant search strings applied individually for each quality indicator. Following the application of inclusion and exclusion criteria, the selection of articles based on content of interest, and searching article references for studies that may have been missed in the initial search, 52 articles remained for analysis (Figure 1). Twelve articles reported success rates for more than one indicator. Twenty-three articles reported bile duct cannulation success rates, 2 articles reported main pancreatic duct cannulation success rates, 18 articles reported precut utilization rates, 15 articles reported common bile duct stone extraction rates, and 11 articles reported biliary stenting rates.

For assessing risk of bias across studies, we selected three biases that commonly affect observational studies—selection, information, and confounding—and evaluated the studies for these biases using the Cochrane Collaboration's summary scale (low, unclear, and high risk of bias). Figure 2 summarizes our risk of bias assessment. For bile and pancreatic duct cannulation, as well as precut utilization, the largest potential bias was confounding bias, with unclear or high risk present in 65, 50, and 69 % of trials, respectively. For common bile

duct stone extraction, selection bias was most prominent, with 60 % of trials having either unclear or high risk of selection bias. For biliary stent placement, selection and information bias were highest, with 36 % of trials having unclear risk.

The cumulative, weighted bile duct cannulation success rate was 89.3% (95% CI 0.866–0.919); pancreatic duct cannulation was 85.0% (95% CI 0.813–0.886); precut utilization rate was 10.5% (95% CI 0.087–0.123); common bile duct stone extraction rate was 88.3 % (95% CI 0.825–0.941); and the rate of successful biliary stenting below the common bile duct bifurcation was 97.5 % (95 % CI 0.967–0.984). Subgroup analysis with meta-regression showed no statistically significant differences between academic and community settings, prospective and retrospective study designs, and trainee participation on success of bile duct cannulation, precut utilization, and common bile duct stone extraction (insufficient observations/variance for pancreatic duct cannulation and biliary stent placement). The results for each quality indicator are tabulated in Tables 1-5 (refs 5-55).

For subgroup analysis by geographical region, there was high geographic region variability across studies. Studies were therefore grouped by continent (Americas, Europe, Asia, and Australia). Success rates were then calculated for each continent and are summarized in Table 6. Overall, continental success rates were in-line with individual cumulative rates, with the exception that common bile duct stone extraction in Europe was lower at 79% and bile duct cannulation in Asia was higher at 98.1%.

## DISCUSSION

This study assessed ERCP intraprocedural performance by examining success rates found in the published literature for the ERCP intraprocedural quality indicators. Of the 52 studies included in our analysis, only one study specifically measured ERCP quality (11); all other studies were designed for other purposes. Because of between-study heterogeneity causing the true effect size to vary from one study to the next, a random-effects model was used. A fixed-effect model, in contrast, assumes that all of the included studies estimate the same effect size. From a weighting perspective, our random-effects model estimated the distribution of the true effects, allowing for a more balanced weighting and a lower risk of larger studies dominating the analysis.

Overall, our analysis reflects that ERCP intraprocedural performance is in good standing. The ACG/ASGE target for bile and main pancreatic duct cannulation is > 85% of ERCP procedures, and our study revealed 89.3 and 85.0% respective success rates for these two indicators. However, we only found two studies that reported a main pancreatic duct cannulation rate; granted, pancreatic duct cannulation is often not the goal of ERCP and thus is harder to assess from a performance perspective. The precut utilization target of < 15% of ERCPs put forth by the taskforce is also being met, with our analysis showing a utilization rate of 10.5%. The common bile duct stone extraction target for standard techniques, including mechanical lithotripsy, is > 90%. Our analysis revealed a success rate of 88.3%. This slightly lower rate may be because five studies in our analysis were limited to stones > 10 mm in size. Successful biliary stent placement had the highest level of performance with a success rate of 97.5 %, exceeding the taskforce's > 80% target. However, for many studies it was unclear whether the reported biliary stent success rate included stent-indicated procedures where biliary cannulation failed. This is an important distinction because once cannulation is achieved and a wire is successfully placed, stent placement is almost always successful.

We did not observe any statistically significant impact on ERCP intraprocedural quality based on health-care setting, study design, or trainee participation. This observation,

however, is tempered by a small number of community-only trials as well as inconsistent reporting of trainee involvement. Similarly, a recent study of US endoscopists found no difference in academic vs. community hospitals across several ERCP intraprocedural quality metrics, namely biliary cannulation and common bile duct stone extraction (56); however, lower success for biliary cannulation and extraction of large stones (> 10 mm) was observed in endoscopists performing < 100 annual ERCPs (although the success rates for each—92 and 91% respectively—were still acceptable).

There are several important limitations of this study. The first is all of the studies included in our analysis, with the exception of one, were designed for purposes other than assessing ERCP intraprocedural quality. This is statistically evident by the high heterogeneity we observed in our meta-analysis calculations ( $I^2$  values > 90%). However, as our focus was extracting success rates for each intraprocedural quality indicator, and our success rates were weighted, the observed heterogeneity between studies becomes less important.

Bias in the studies reviewed is also an important limitation. Unfortunately, there is neither a validated nor accepted tool for assessing bias in observational studies. We assessed bias based on three common biases found in observational studies: selection bias, information bias, and confounding. Suspicion for selection bias—a sample population that may not be representative of the population of interest—arose when study populations were limited to certain demographics, disease states, or ERCP indications. For information bias, retrospective designs were more often implicated than prospective designs. Confounding bias was suspected when factors that may affect the intraprocedural quality indicator of interest were not clearly accounted for in the statistical methods or mentioned in the text as a limitation. Examples of confounding factors include: uncertain papilla status; unclear endoscopist experience level or volume of procedures; unclear trainee participation; and unknown stone size distribution.

Another limitation is that any meta-analysis or systematic review has the potential to miss important studies that may impact the statistical results. We believe our analysis overcomes this limitation by the large number of procedures (> 30,000) that were included. It is unlikely a missed study or even several missed studies would significantly impact the cumulative, weighted success rates for each of the quality indicators.

Finally, the greatest limitation of this analysis is that it was confined to the published literature only and therefore only reflects the ERCP performance of institutions, primarily academic, that are conducting clinical research. It is very difficult to generalize this performance assessment to the broader ERCP community.

Despite these limitations, we believe the results from this analysis can be used to help inform and guide future ACG/ASGE taskforce recommendations related to ERCP intraprocedural quality. For future studies on ERCP intraprocedural quality, when reporting success rates for the intraprocedural quality indicators, it will be important to document the health-care setting where the procedure took place in studies that involve both academic and community hospitals (this was not consistently documented in the studies examined for this review that involved both settings). Because partnerships, including research partnerships between academic and community hospitals, are occurring at an increased rate, this differentiation will help us better assess the current state of practice in each respective setting. It will also be important to improve reporting on trainee involvement. In our analysis, trainee involvement was reported at a very low rate: only 32% (13/40) of studies adequately described the extent of trainee involvement.

Regarding the actual intraprocedural quality indicators, consideration could be given to encouraging improved reporting of main pancreatic duct cannulation success rates (when

pancreatic duct cannulation is indicated). One new indicator that may be of interest, both from a safety as well as a training perspective, would be time to ductal cannulation. Moreover, the two targets that could be potentially revised would be precut utilization and biliary stenting. Our analysis shows performance is exceeding these targets, especially for biliary stent placement. However, it should be noted that most of the studies reviewed for biliary stenting focused on placement of self-expanding metallic stents for both benign and malignant conditions. Proper placement of a self-expanding metallic stents is sometimes more difficult than placement of a plastic stent, but the overall success rate was still very high. As such, this analysis may not accurately represent the success rate for placement of a plastic biliary stent, especially in a community setting. Despite this, it could be reasonable to consider increasing the biliary stenting target from > 80 to > 90% and consider decreasing the precut target from 15 to 10%. But again, the important caveat, as said earlier regarding these target revision recommendations, is the majority of hospitals in this analysis were academic. To truly consider changing targets, it is imperative to gather more data from community-based centers. Moreover, with regard to biliary stenting, it will be important in future studies to clearly state whether the reported success rate includes or does not include cases where biliary cannulation fails.

It would also be beneficial to report stone size more consistently. Five of the fifteen studies included in our common bile duct stone extraction analysis included only stones > 10 mm. The remaining studies presumably included stones of all sizes, but the size distribution is not known. Stone size (< 10 or > 10 mm) is certainly important when grading procedural difficulty, and success rates for stone extraction should be similarly divided for small and large stones.

In conclusion, health-care quality is something that will undoubtedly demand our attention in the years ahead. We hope this analysis serves as a starting point for encouraging the measurement and reporting of ERCP intraprocedural quality on a consistent and reliable basis. Currently, there are two gastroenterology registries available for reporting quality metrics: the American Gastroenterological Association's (AGA) Digestive Health Outcomes Registry and the GI Quality Improvement Consortium. The capacity for ERCP reporting within these registries will hopefully become available in the near future. Increased reporting will help us continue to set appropriate quality metrics and benchmarks as we move forward, and allow us to quickly respond to new data and technology. Increased reporting will also help us refine our metrics as a whole—the importance of which cannot be understated, as we enter an age when reimbursement in every facet of the health-care landscape will be increasingly tied to performance and the quality of care that we deliver.

## Acknowledgments

**Financial support** : Dr DeBenedet's contribution was supported by grant 5T32DK062708-09 from the National Institutes of Health. Dr Elmunzer's contribution was supported by grant UL1RR024986 from the National Center for Research Resources. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Center for Research Resources. Dr Schoenfeld's contribution was supported by grant K24-1K24DK084208 from the National Institutes of Health.

## References

1. Dorn S. Gastroenterology in a new era of accountability: Part 1. An overview of performance measurement. *Clin Gastro Hep*. 2011; 9:563–6.
2. Young, PL.; Olsen, L.; McGinnis, JM. Value in health care: accounting for cost, quality, safety, outcomes, and innovation. National Academies Press; Washinton, DC: 2010.
3. Baron TH, Petersen BT, Mergener K, et al. Quality indicators for endoscopic retrograde cholangiopancreatography. *Am J Gastroenterol*. 2006; 101:892–7. [PubMed: 16635233]

4. Baron TH, Petersen BT, Mergener K, et al. Quality indicators for endoscopic retrograde cholangiopancreatography. *Gastrointest Endosc.* 2006; 63(Suppl 4):S29–34. [PubMed: 16564909]
5. Adler DG, Verma D, Hilden K, et al. Dye-free wire-guided cannulation of the biliary tree during ERCP is associated with high success and low complication rates: outcomes in a single operator experience of 822 cases. *J Clin Gastroenterol.* 2010; 44:e57–62. [PubMed: 19636260]
6. Ayoubi M, Sansoe G, Leone N, et al. Comparison between needle-knife fistulotomy and standard in ERCP. *World J Gastrointest Endosc.* 2012; 4:398–404. [PubMed: 23125897]
7. Atamanalp S, Yildirman M, Kantarci A. Endoscopic retrograde cholangiopancreatography (ERCP): outcomes of 3136 cases over 10 years. *Turk J Med Sci.* 2011; 41:615–21.
8. Belverde B, Frattaroli S, Carbone A, et al. Double guidewire technique for ERCP in difficult bile cannulation: experience with 121 cases. *Ann Ital Chir.* 2012; 83:391–3. [PubMed: 23064299]
9. Cennamo V, Fuccio L, Repici A, et al. Timing of precut procedure does not influence success rate and complications of ERCP procedure: a prospective randomized comparative study. *Gastrointest Endosc.* 2009; 69(3 Part 1):473–9. [PubMed: 19231488]
10. Chatterjee S, Rees C, Dwarakanath AD, et al. Endoscopic retrograde cholangio-pancreatography practice in district general hospitals in North East England: a Northern Regional Endoscopy Group (NREG) study. *J R Coll Physicians Edinb.* 2011; 41:109–13. [PubMed: 21677912]
11. Colton JB, Curran CC. Quality indicators, including complications, of ERCP in a community setting: a prospective study. *Gastrointest Endosc.* 2009; 70:457–67. [PubMed: 19482278]
12. Enochsson L, Swahn F, Arnelo U, et al. Nationwide, population-based data from 11,074 ERCP procedures from the Swedish Registry for Gallstone Surgery and ERCP. *Gastrointest Endosc.* 2010; 72:1175–84. 1184. [PubMed: 20970787]
13. García-Cano J, González-Martín JA. Bile duct cannulation: success rates for various ERCP techniques and devices at a single institution. *Acta Gastroenterol Belg.* 2006; 69:261–7. [PubMed: 17168121]
14. Glomsaker T, Søreide K, Hoff G, et al. Contemporary use of endoscopic retrograde cholangiopancreatography (ERCP): a Norwegian prospective, multicenter study. *Scand J Gastroenterol.* 2011; 46:1144–51. [PubMed: 21692712]
15. Imazu H, Ikeda K, Kakutani H, et al. A pilot study of the novel offset-tip papillotome for selective biliary cannulation in ERCP. *Minim Invasive Ther Allied Technol.* 2012; 21:335–41. [PubMed: 22098440]
16. Khatibian M, Sotoudehmanesh R, Ali-Asgari A, et al. Needle-knife fistulotomy vs. standard method for cannulation of common bile duct: a randomized controlled trial. *Arch Iran Med.* 2008; 11:16–20. [PubMed: 18154417]
17. Kouklakis G, Gatopoulou A, Lirantzopoulos N, et al. Evaluation of guide wire cannulation technique in elderly patients with choledocholithiasis. *J Gastroint Liver Dis.* 2009; 18(2):185–8.
18. Laohavichitra K, Akaraviputh T, Methasate A, et al. Comparison of early pre-cutting vs standard technique for biliary cannulation in endoscopic retrograde cholangiopancreatography: a personal experience. *World J Gastroenterol.* 2007; 13:3734–7. [PubMed: 17659735]
19. Li QY, Pan L, Ling Q, et al. Single-operator wire-guided cannulation technique enables easier cannulation of endoscopic retrograde cholangiopancreatography. *Dig Dis Sci.* 2012; 57:3293–8. [PubMed: 22736016]
20. Mariani A, Giussani A, Di Leo M, et al. Guidewire biliary cannulation does not reduce post-ERCP pancreatitis compared with the contrast injection technique in low-risk and high-risk patients. *Gastrointest Endosc.* 2012; 75:339–46. [PubMed: 22075192]
21. Nakai Y, Isayama H, Tsujino T, et al. Impact of introduction of wire-guided cannulation in therapeutic biliary endoscopic retrograde cholangiopancreatography. *J Gastroenterol Hepatol.* 2011; 26:1552–8. [PubMed: 21615792]
22. Panteris V, Vezakis A, Filippou G, et al. Influence of juxtaepapillary diverticula on the success or difficulty of cannulation and complication rate. *Gastrointest Endosc.* 2008; 68:903–10. [PubMed: 18635174]
23. Testoni PA, Mariani A, Giussani A, et al. Risk factors for post-ERCP pancreatitis in high- and low-volume centers and among expert and non-expert operators: a prospective multicenter study. *Am J Gastroenterol.* 2010; 105:1753–61. [PubMed: 20372116]

24. Testoni PA, Giussani A, Vailati C, et al. Precut sphincterotomy, repeated cannulation and post-ERCP pancreatitis in patients with bile duct stone disease. *Dig Liver Dis.* 2011; 43:792–6. [PubMed: 21733768]
25. Trifan A, Sfarti C, Cretu M, et al. Guide-wire vs. conventional contrast cannulation of the common bile duct for the prevention of post-ERCP pancreatitis in patients with choledocholithiasis. *J Gastrointestin Liver Dis.* 2011; 20:149–52. [PubMed: 21725511]
26. Williams EJ, Taylor S, Fairclough P, et al. Are we meeting the standards set for endoscopy? Results of a large-scale prospective survey of endoscopic retrograde cholangiopancreatograph practice. *Gut.* 2007; 56:821–9. [PubMed: 17145737]
27. Xinopoulos D, Bassioulas SP, Kypreos D, et al. Pancreatic duct guidewire placement for biliary cannulation in a single-session therapeutic ERCP. *World J Gastroenterol.* 2011; 17:1989–95. [PubMed: 21528077]
28. Donnellan F, Zeb F, Courtney G, et al. Suprapapillary needleknife fistulotomy: a safe and effective method for accessing the biliary system. *Surg Endosc.* 2010; 24:1937–40. [PubMed: 20135176]
29. Fukatsu H, Kawamoto H, Kato H, et al. Evaluation of needle-knife precut papillotomy after unsuccessful biliary cannulation, especially with regard to postoperative anatomic factors. *Surg Endosc.* 2008; 22:717–23. [PubMed: 17704885]
30. Kapral C, Duller C, Wewalka F, et al. Case volume and outcome of endoscopic retrograde cholangiopancreatography: results of a nationwide Austrian benchmarking project. *Endoscopy.* 2008; 40:625–30. [PubMed: 18680074]
31. Katsinelos P, Gkagkalis S, Chatzimavroudis G, et al. Comparison of three types of precut technique to achieve common bile duct cannulation: a retrospective analysis of 274 cases. *Dig Dis Sci.* 2012; 57(12):3286–92. [PubMed: 22714730]
32. Robison LS, Varadarajulu S, Wilcox CM. Safety and success of precut biliary sphincterotomy: is it linked to experience or expertise? *World J Gastroenterol.* 2007; 13:2183–6. [PubMed: 17465498]
33. Siddiqui AR, Niaz SK. Needle knife papillotomy for cannulating difficult papilla; two years experience. *J Pak Med Assoc.* 2008; 58:195–7. [PubMed: 18655429]
34. Attasaranya S, Cheon YK, Vittal H, et al. Large-diameter biliary orifice balloon dilation to aid in endoscopic bile duct stone removal: a multicenter series. *Gastrointest Endosc.* 2008; 67:1046–52. [PubMed: 18178208]
35. Ito Y, Tsujino T, Togawa O, et al. Endoscopic papillary balloon dilation for the management of bile duct stones in patients 85 years of age and older. *Gastrointest Endosc.* 2008; 68:477–82. [PubMed: 18760175]
36. Itoi T, Itokawa F, Sofuni A, et al. Endoscopic sphincterotomy combined with large balloon dilation can reduce the procedure time and fluoroscopy time for removal of large bile duct stones. *Am J Gastroenterol.* 2009; 104:560–5. [PubMed: 19174779]
37. 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; 70:915–22. [PubMed: 19647241]
38. Kim HJ, Choi HS, Park JH, et al. Factors influencing the technical difficulty of endoscopic clearance of bile duct stones. *Gastrointest Endosc.* 2007; 66:1154–60. [PubMed: 17945223]
39. Kochhar R, Dutta U, Shukla R, et al. Sequential endoscopic papillary balloon dilatation following limited sphincterotomy for common bile duct stones. *Dig Dis Sci.* 2009; 54:1578–81. [PubMed: 19005760]
40. Lee SH, Park JK, Yoon WJ, et al. How to predict the outcome of endoscopic mechanical lithotripsy in patients with difficult bile duct stones? *Scand J Gastroenterol.* 2007; 42:1006–10. [PubMed: 17613932]
41. Masci E, Minoli G, Rossi M, et al. Prospective multicenter quality assessment of endotherapy of biliary stones: does center volume matter? *Endoscopy.* 2007; 39:1076–81. [PubMed: 18072060]
42. Minami A, Hirose S, Nomoto T, et al. Small sphincterotomy combined with papillary dilation with large balloon permits retrieval of large stones without mechanical lithotripsy. *World J Gastroenterol.* 2007; 13:2179–82. [PubMed: 17465497]
43. Siiki A, Tamminen A, Tomminen T, et al. ERCP procedures in a Finnish community hospital: a retrospective analysis of 1207 cases. *Scand J Surg.* 2012; 101:45–50. [PubMed: 22414468]



44. Tsujino T, Yoshida H, Isayama H, et al. Endoscopic papillary balloon dilation for bile duct stone removal in patients 60 years old or younger. *J Gastroenterol.* 2010; 45:1072–9. [PubMed: 20467759]
45. Youn YH, Lim HC, Jahng JH, et al. The increase in balloon size to over 15mm does not affect the development of pancreatitis after endoscopic papillary large balloon dilatation for bile duct stone removal. *Dig Dis Sci.* 2011; 56:1572–7. [PubMed: 20945093]
46. Banerjee N, Hilden K, Baron TH, et al. Endoscopic biliary sphincterotomy is not required for transpapillary SEMS placement for biliary obstruction. *Dig Dis Sci.* 2011; 56:591–5. [PubMed: 20632105]
47. Behm B, Brock A, Clarke BW, et al. Partially covered self-expandable metallic stents for benign biliary strictures due to chronic pancreatitis. *Endoscopy.* 2009; 41:547–51. [PubMed: 19533560]
48. Gómez-Oliva C, Guarner-Argente C, Concepción M, et al. Partially covered self-expanding metal stent for unresectable malignant extrahepatic biliary obstruction: results of a large prospective series. *Surg Endosc.* 2012; 26:222–9. [PubMed: 21858574]
49. Han YM, Kwak HS, Jin GY, et al. Treatment of malignant biliary obstruction with a PTFE-covered self-expandable nitinol stent. *Korean J Radiol.* 2007; 8:410–7. [PubMed: 17923784]
50. Kahaleh M, Behm B, Clarke BW, et al. Temporary placement of covered self-expandable metal stents in benign biliary strictures: a new paradigm? with video. *Gastrointest Endosc.* 2008; 67:446–54. [PubMed: 18294506]
51. Siddiqui AA, Mehendiratta V, Loren D, et al. Self-expanding metal stents (SEMS) for preoperative biliary decompression in patients with resectable and borderline-resectable pancreatic cancer: outcomes in 241 patients. *Dig Dis Sci.* 2012; 58:44–50.
52. Tsuchiya T, Itoi T, Gotoda T, et al. A multicenter prospective study of the short-term outcome of a newly developed partially covered self-expandable metallic biliary stent (WallFlex<sup>®</sup>). *Dig Dis Sci.* 2011; 56:1889–95. [PubMed: 21298481]
53. Yamaguchi T, Ishihara T, Seza K, et al. Long-term outcome of endoscopic metallic stenting for benign biliary stenosis associated with chronic pancreatitis. *World J Gastroenterol.* 2006; 12:426–30. [PubMed: 16489643]
54. Yoon WJ, Lee JK, Lee KH, et al. A comparison of covered and uncovered Wallstents for the management of distal malignant biliary obstruction. *Gastrointest Endosc.* 2006; 63:996–1000. [PubMed: 16733115]
55. Yoon WJ, Ryu JK, Yang KY, et al. A comparison of metal and plastic stents for the relief of jaundice in unresectable malignant biliary obstruction in Korea: an emphasis on cost-effectiveness in a country with a low ERCP cost. *Gastrointest Endosc.* 2009; 70:284–9. [PubMed: 19539921]
56. Cotton PB, Romagnuolo J, Faigel DO, et al. The ERCP quality network: a pilot study of benchmarking practice and performance. *Am J Med Qual.* 2012; 28:256–60. [PubMed: 22930708]

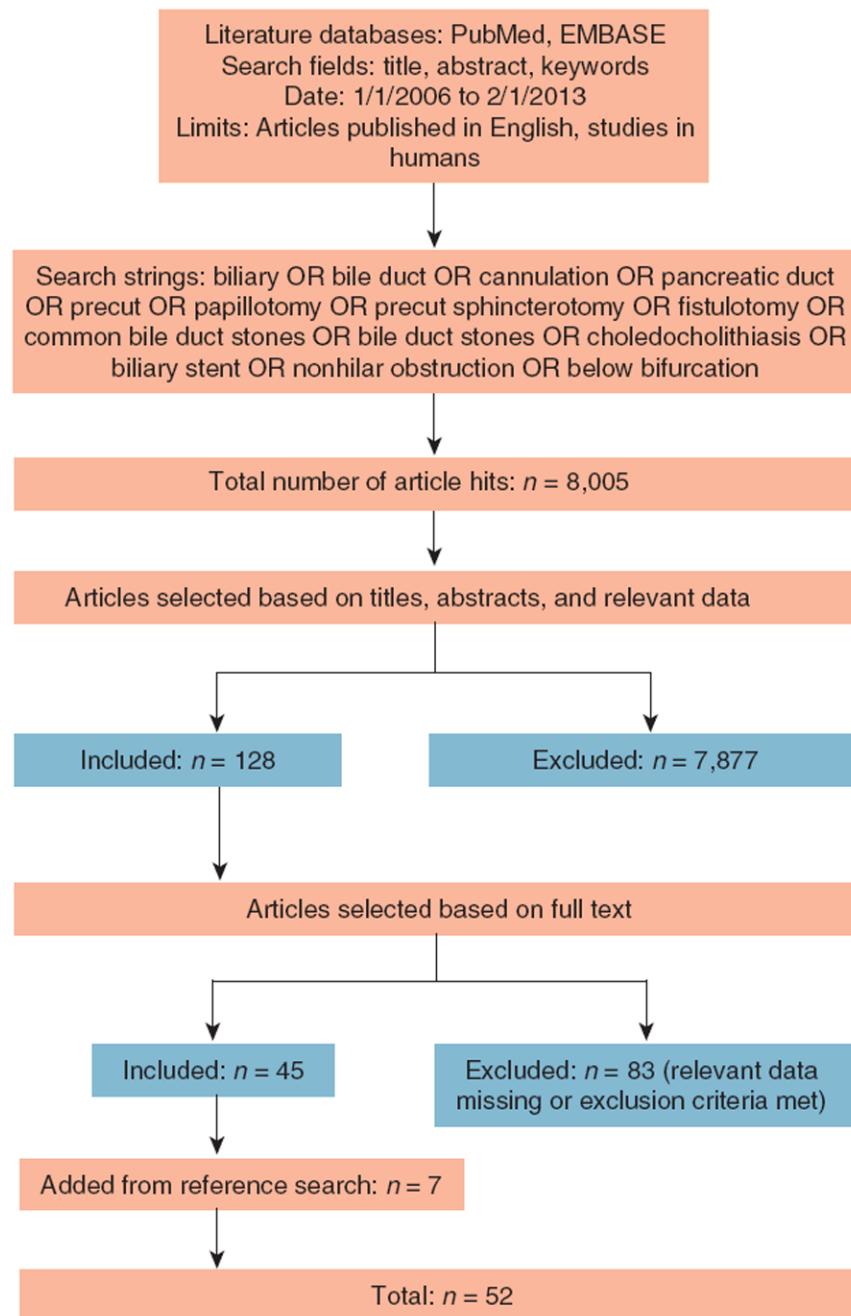
### Study Highlights

#### WHAT IS CURRENT KNOWLEDGE

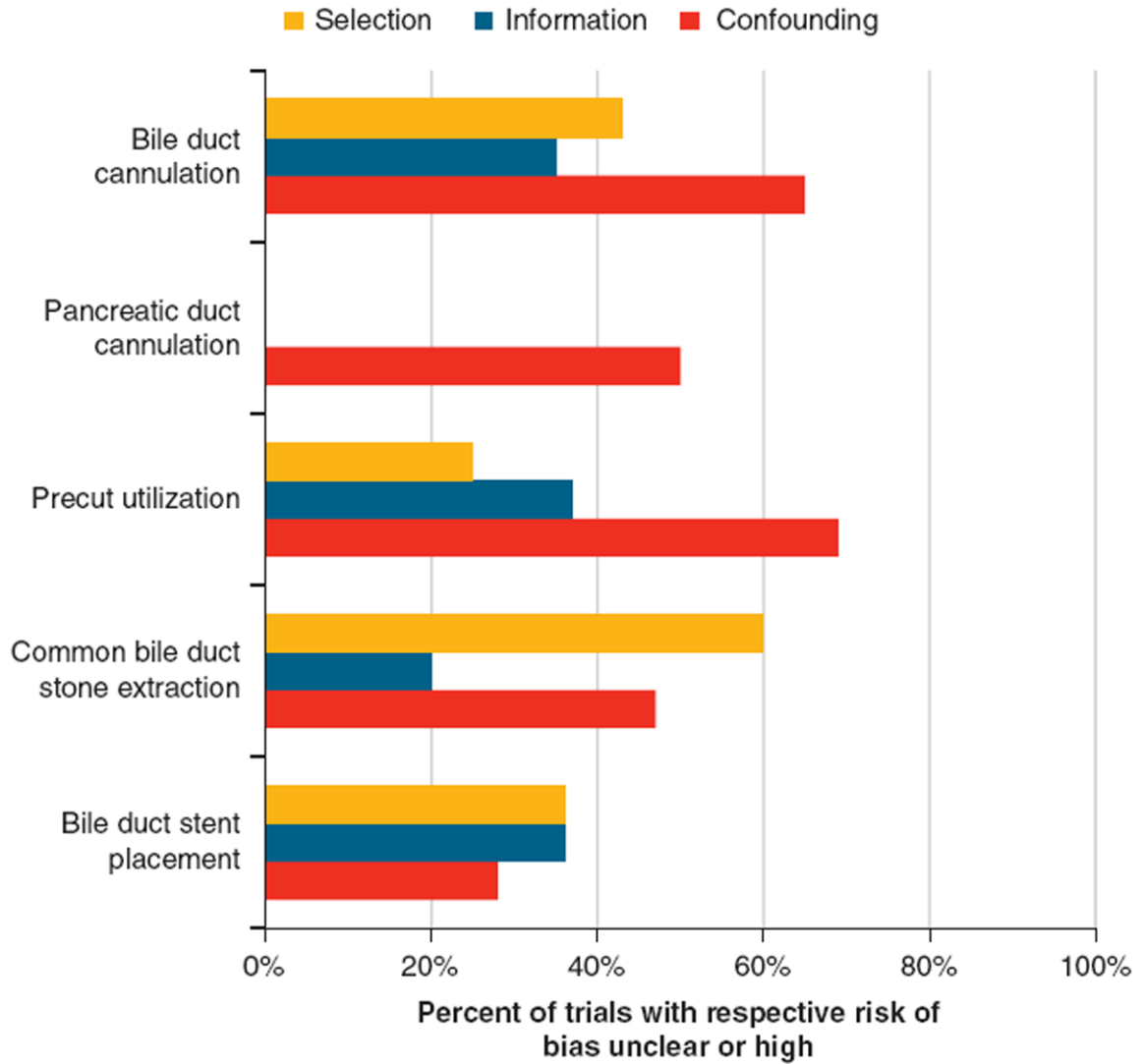
- ✓ Quality in endoscopy is increasingly becoming an important topic of inquiry and research.
- ✓ In 2006, experts defined quality indicators for endoscopic procedures, including endoscopic retrograde cholangiopancreatography (ERCP).
- ✓ Benchmarks for each indicator were also defined.

#### WHAT IS NEW HERE

- ✓ We are meeting or exceeding benchmarks for ERCP intra procedural quality; however, this performance is mainly reflective of academic practice.
- ✓ Modifying the precut utilization and biliary stent benchmarks could be considered in the future.
- ✓ Health-care setting (academic vs. community), study design (retrospective vs. prospective), and trainee involvement had no impact on ERCP intraprocedural quality.



**Figure 1.**  
Article selection methodology.



**Figure 2.**

Risk of bias assessment. Suspicion for selection bias arose when study populations were limited to certain demographics, disease states, or endoscopic retrograde cholangiopancreatography indications. Information bias arose predominantly in retrospectively designed studies. Confounding bias was suspected when factors that may affect the intraprocedural quality indicator of interest were not clearly accounted for.

Table 1

## Bile duct cannulation success

Study	Setting	Design	Trainee	Success rate	s.e.	95 % CI	% Weight
Adler <i>et al.</i> (5)	A	R	NA	801/822=0.974	0.006	0.964-0.985	4.66
Atamanalp <i>et al.</i> (7)	A	R	N	2,965/3,664=0.809	0.006	0.800-0.822	4.65
Ayoubi <i>et al.</i> (6)	C	R	NA	60/85=0.706	0.049	0.609-0.803	2.87
Belverde <i>et al.</i> (8)	A	P	Y	1,214/1,335=0.909	0.009	0.853-0.888	4.60
Cennamo <i>et al.</i> (9)	A	P	NA	78/110=0.709	0.043	0.624-0.794	3.16
Chatterjee <i>et al.</i> (10)	A/C	R	Y	420/481=0.873	0.015	0.843-0.903	4.43
Colton <i>et al.</i> (11)	C	P	N	692/764=0.906	0.011	0.885-0.926	4.56
Enochsson <i>et al.</i> (12)	A/C	P	Y	10,188/11,074=0.920	0.003	0.915-0.925	4.69
Garcia-Cano <i>et al.</i> (13)	A	P	Y	168/199=0.844	0.026	0.794-0.895	4.01
Glomsaker <i>et al.</i> (14)	A/C	P	NA	3,278/3,809=0.861	0.006	0.850-0.871	4.66
Imazu <i>et al.</i> (15)	A	R	NA	77/85=0.906	0.032	0.844-0.968	3.73
Khatibian <i>et al.</i> (16)	A	P	NA	100/106=0.943	0.022	0.899-0.987	4.15
Kouklakis <i>et al.</i> (17)	A	R	NA	201/217=0.926	0.018	0.891-0.961	4.34
Laohavichitra <i>et al.</i> (18)	A	R	NA	287/293=0.980	0.008	0.963-0.996	4.61
Li <i>et al.</i> (19)	A	P	Y	460/465=0.989	0.005	0.980-0.999	4.67
Mariani <i>et al.</i> (20)	A	P	Y	1,149/1,249=0.920	0.008	0.905-0.935	4.63
Nakai <i>et al.</i> (21)	A	P	Y	493/500=0.986	0.005	0.976-0.996	4.66
Panteris <i>et al.</i> (22)	C	R	N	570/601=0.948	0.018	0.891-0.961	4.34
Testoni <i>et al.</i> (23)	A/C	P	Y	3,213/3,635=0.884	0.005	0.873-0.894	4.66
Testoni <i>et al.</i> (24)	A	R	Y	1,717/2,004=0.857	0.008	0.841-0.872	4.62
Trifan <i>et al.</i> (25)	A	R	NA	115/128=0.898	0.027	0.846-0.951	3.96
Williams <i>et al.</i> (26)	A/C	P	Y	4,089/5,264=0.777	0.006	0.766-0.788	4.66
Xinopoulos <i>et al.</i> (27)	A	R	NA	2,202/2,332=0.944	0.005	0.935-0.954	4.67
Cumulative weighted success rate				0.893		0.866-0.919	100.0

Coef.	95 % CI	P value
Setting	-0.063 to 0.157	0.377

Design	-0.008	-0.074 to 0.058	0.803
Trainee	0.036	-0.077 to 0.149	0.497

Heterogeneity  $\tau^2=1.756.02$  (d.f.=22),  $P=0.000$ ;  $\hat{P}=98.7\%$ ; estimate of between-study variance  $\tau^2=0.004$ ; test of ES=0:  $z=66.39$ ,  $P=0.000$

#### Metaregression subgroup analysis

CI, confidence interval; Coef., coefficient; ES, effect size; NA, not applicable.

Table 2

## Pancreatic duct cannulation success

Study	Setting	Design	Trainee	Success rate	s.e.	95% CI	% Weight
Colton <i>et al.</i> (11)	C	P	N	81/93=0.871	0.035	0.803-0.939	28.53
Glomsaker <i>et al.</i> (14)	A/C	P	Y	233/277=0.841	0.022	0.798-0.884	71.47
Cumulative weighted success rate				0.850		0.813-0.886	100.0

Heterogeneity  $\tau^2=0.53$  (d.f.=1),  $P=0.468$ ;  $\hat{I}^2=0.0\%$ ; test of ES=0:  $z=45.76$ ,  $P=0.000$

Metaregression subgroup analysis

Insufficient observations for analysis

CI, confidence interval; ES, effect size; NA, not applicable.

Table 3

## Pre-cut utilization

Study	Setting	Design	Trainee	Utilization rate	s.e.	95% CI	% Weight
Atamanalp <i>et al.</i> (7)	A	R	N	465/3,664=0.127	0.005	0.116–0.137	6.19
Ayoubi <i>et al.</i> (6)	C	R	NA	25/85=0.294	0.049	0.197–0.391	2.28
Cennamo <i>et al.</i> (9)	A	P	NA	32/110=0.291	0.043	0.206–0.375	2.68
Chatterjee <i>et al.</i> (10)	A/C	R	NA	4/481=0.008	0.004	0.000–0.016	6.25
Colton <i>et al.</i> (11)	C	P	N	26/805=0.032	0.006	0.020–0.045	6.15
Donnellan <i>et al.</i> (28)	A	R	Y	352/2,603=0.135	0.006	0.122–0.148	6.12
Enochsson <i>et al.</i> (12)	A/C	P	Y	858/1,1074=0.077	0.002	0.072–0.082	6.30
Fukatsu <i>et al.</i> (29)	A	R	NA	80/501=0.160	0.016	0.127–0.191	5.30
Garcia-Cano <i>et al.</i> (13)	A	P	Y	23/199=0.116	0.022	0.071–0.160	4.61
Glomsaker <i>et al.</i> (14)	A/C	P	NA	190/3,809=0.050	0.003	0.042–0.057	6.27
Katsinelos <i>et al.</i> (31)	A	R	N	283/2,903=0.097	0.006	0.087–0.108	6.19
Mariani <i>et al.</i> (20)	A	P	Y	100/1,249=0.080	0.007	0.065–0.095	6.07
Panteris <i>et al.</i> (22)	C	R	N	164/601=0.223	0.018	0.237–0.308	5.11
Robison <i>et al.</i> (32)	A	R	NA	150/2,939=0.051	0.004	0.043–0.059	6.25
Siddiqui <i>et al.</i> (33)	C	R	N	59/515=0.115	0.014	0.087–0.142	5.54
Testoni <i>et al.</i> (23)	A/C	P	Y	308/3,635=0.085	0.004	0.076–0.094	6.23
Testoni <i>et al.</i> (24)	A	R	Y	161/2,004=0.080	0.006	0.068–0.092	6.16
Williams <i>et al.</i> (26)	A/C	P	Y	465/5,264=0.088	0.004	0.081–0.096	6.26
Cumulative weighted utilization rate				0.105		0.087–0.123	100.0

Coef.	95% CI	P value	
Setting	-0.047	-0.161 to 0.067	0.383
Design	-0.032	-0.113 to 0.049	0.415
Trainee	-0.032	-0.107 to 0.0	0.359

Heterogeneity  $I^2=794.94$  (d.f.=17),  $P=0.000$ ;  $I^2=97.9\%$ ; estimate of between-study variance  $\tau^2=0.001$ ; test of ES=0:  $Z=11.25$ ,  $P=0.000$

## Metaregression subgroup analysis

CI, confidence interval; Coef., coefficient; ES, effect size; NA, not applicable.



Table 4

Common bile duct stone extraction success

Study	Setting	Design	Trainee	Success rate	s.e.	95% CI	% Weight
Atasaryana <i>et al.</i> (34)	A	R	Y	102/107=0.953	0.020	0.913-0.993	6.68
Colton <i>et al.</i> (11)	C	P	N	80/92=0.870	0.035	0.800-0.938	6.29
Enochsson <i>et al.</i> (12)	A/C	P	Y	3,140/4,074=0.771	0.006	0.757-0.784	6.89
Ito <i>et al.</i> (35)	A	R	NA	273/406=0.672	0.011	0.650-0.694	6.83
Itoi <i>et al.</i> (36)	A	R	NA	92/101=0.911	0.010	0.892-0.930	6.85
Jeong <i>et al.</i> (37)	A	R	NA	33/38=0.868	0.026	0.818-0.919	6.55
Kapral <i>et al.</i> (30)	A	P	Y	1,124/1,140=0.986	0.003	0.979-0.993	6.90
Kim <i>et al.</i> (38)	A	P	NA	96/102=0.941	0.023	0.896-0.987	6.62
Kochhar <i>et al.</i> (39)	A	P	NA	62/74=0.838	0.032	0.776-0.900	6.39
Lee <i>et al.</i> (40)	A	R	NA	102/134=0.761	0.037	0.689-0.833	6.24
Masci <i>et al.</i> (41)	A/C	P	Y	504/580=0.869	0.014	0.842-0.896	6.80
Minami <i>et al.</i> (42)	A	P	NA	87/88=0.989	0.011	0.966-1.011	6.83
Silki <i>et al.</i> (43)	C	R	N	243/266=0.913	0.017	0.880-0.947	6.74
Tsujino <i>et al.</i> (44)	A	P	NA	304/311=0.977	0.008	0.961-0.994	6.86
Youn <i>et al.</i> (45)	A	R	NA	93/101=0.921	0.027	0.868-0.973	6.53
Cumulative weighted success rate				0.883		0.825-0.941	100.0

Coef.	95% CI	P-value
Setting	0.010 -1.123 to 0.143	0.875
Design	0.049 -0.055 to 0.152	0.331
Trainee	0.001 -0.211 to 0.214	0.986

Heterogeneity  $\tau^2=1,486.37$  (d.f.=14),  $P=0.000$ ;  $\hat{\tau}^2=99.1\%$ ; estimate of between-study variance  $\tau^2=0.0127$ ; test of ES=0:  $z=29.86$ ,  $P=0.000$

Metaregression subgroup analysis

CI, confidence interval; Coef., coefficient; ES, effect size; NA, not applicable.

Table 5

## Biliary stent placement success

Study	Setting	Design	Trainee	Success rate	s.e.	95% CI	% Weight
Banejee <i>et al.</i> (46)	A	R	NA	104/104=1.000	0	NA	NA
Behm <i>et al.</i> (47)	A	P	NA	19/20=0.950	0.049	0.854-1.046	0.790
Gómez-Olivía <i>et al.</i> (48)	A	P	NA	191/199=0.960	0.014	0.933-0.987	9.68
Han <i>et al.</i> (49)	A	P	NA	37/37=1.000	0	NA	NA
Kahaleh <i>et al.</i> (50)	A	P	NA	79/79=1.000	0	NA	NA
Kapral <i>et al.</i> (30)	A	P	Y	1,034/1,058=0.977	0.005	0.968-0.986	89.53
Siddiqui <i>et al.</i> (33)	A	R	NA	241/241=1.000	0	NA	NA
Tsuchiya <i>et al.</i> (52)	A	P	NA	52/52=1.000	0	NA	NA
Yamaguchi <i>et al.</i> (53)	A	P	NA	8/8=1.000	0	NA	NA
Yoon <i>et al.</i> (54)	A	R	NA	77/77=1.000	0	NA	NA
Yoon <i>et al.</i> (55)	A	R	NA	112/112=1.000	0	NA	NA
Cumulative weighted success rate				0.975		0.967-0.984	100.0

Heterogeneity  $\chi^2=1.70$  (d.f.=2),  $P=0.427$ ;  $I^2=0.0\%$ ; estimate of between-study variance  $\tau^2=0.000$ ; test of ES=0:  $z=225.19$ ,  $P=0.000$

Metaregression subgroup analysis

Insufficient variance for analysis

CI, confidence interval; ES, effect size; NA, not applicable.

**Table 6**

Success rates by geographic region

Intraoperative quality indicator	Americas	Europe	Asia	Australia
Bile duct cannulation	1,913/2,067=0.925	31,307/35,812=0.874	1,317/1,343=0.981	NA
Pancreatic duct cannulation	81/93=0.871	233/277=0.841	NA	NA
Precut utilization	180/4,225=0.042	3,426/37,200=0.092	139/1,016=0.137	NA
Common bile duct stone extraction	182/199=0.915	3,887/4,920=0.790	1,142/1,355=0.843	1,124/1,140=0.986
Biliary stent placement	443/444=0.998	191/199=0.960	286/286=1.000	1,034/1,058=0.977

NA, not applicable.