NIH Public Access
Author Manuscript
Med Care. Author manuscript; available in PMC 2014 December 01.

Published in final edited form as:
Med Care. 2013 December ; 51(12): . doi:10.1097/MLR.0b013e3182a502dc.

Lower Provider Volume is Associated with Higher Failure Rates for Endoscopic Retrograde Cholangiopancreatography

Gregory A. Coté, MD, MS[1], **Timothy D. Imler, MD**[1,5], **Huiping Xu, PhD**[3], **Evgenia Teal, MS**[5], **Dustin D. French, PhD**[2], **Thomas F. Imperiale, MD**[1,5,6], **Marc B. Rosenman, MD**[4,5], **Jeffery Wilson, PhD**[7], **Siu L. Hui, PhD**[3,5], and **Stuart Sherman, MD**[1]

[1] Division of Gastroenterology & Hepatology, Indiana University School of Medicine, Indianapolis, IN

[2] Department of Ophthalmology and Center for Healthcare Studies, Northwestern University Feinberg School of Medicine and VA Center of Excellence, Chicago, IL

[3] Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN

[4] Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

[5] Regenstrief Institute, Inc., Indianapolis, IN

[6] Center of Excellence for Implementation of Evidence-based Practice, Roudebush VA Medical Center, Indianapolis, IN

[7] Department of Geography, Indiana University-Purdue University Indianapolis, Indianapolis, IN

Abstract

Background—Among physicians who perform endoscopic retrograde cholangiopancreatography (ERCP), the relationship between procedure volume and outcome is unknown.

Objective—Quantify the ERCP volume-outcome relationship by measuring provider-specific failure rates, hospitalization rates and other quality measures.

Research Design—Retrospective Cohort

Subjects—16,968 ERCPs performed by 130 physicians between 2001-2011, identified in the Indiana Network for Patient Care (INPC)

Measures—Physicians were classified by their average annual INPC volume and stratified into low (<25/year) and high (≥25/year). Outcomes included failed procedures, defined as repeat ERCP, percutaneous transhepatic cholangiography or surgical exploration of the bile duct ≤ 7 days after the index procedure, hospitalization rates, and 30-day mortality.

Correspondence: Gregory A. Coté, MD, MS Indianapolis University School of Medicine University Hospital 550 N. University Blvd., UH 1634 Indianapolis, IN 46202 gcote@iu.edu.

Conflict of Interest: None of the authors have any conflicts of interest to disclose involving the work under consideration for publication.

Author Contributions: Gregory Coté, Siu Hui, and Huiping Xu had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study design (Coté, Imler), Data collection (Coté, Teal, Imler, Rosenman), Data analysis (Xu, Hui, Wilson), Manuscript drafting (Coté), Critical editing (Coté, Imler, Sherman, Rosenman, French, Imperiale)

Results—Among 15,514 index ERCPs, there were 1,163 (7.5%) failures; the failure rate was higher among low (9.5%) compared to high volume (5.7%) providers (p<0.001). A second ERCP within 7 days (a subgroup of failure rate) occurred more frequently when the original ERCP was performed by a low (4.1%) versus a high volume physician (2.3%, p=0.013). Patients were more frequently hospitalized within 24 hours when the ERCP was performed by a low (28.3%) vs. high volume physician (14.8%, p=0.002). Mortality within 30 days was similar (low – 1.9%, high – 1.9%). Among low volume physicians and after adjusting, the odds of having a failed procedure decreased 3.3% (95% CI 1.6-5.0%, p<0.001) with each additional ERCP performed per year.

Conclusions—Lower provider volume is associated with higher failure rate for ERCP, and greater need for post-procedure hospitalization.

Keywords

ERCP quality; outcomes; gastroenterology

Introduction

There is an increasing emphasis on improving quality of care by implementing minimum volume standards for high-risk procedures. An inverse relationship between provider or facility volume and outcomes has been established for a variety of procedures including upper endoscopy[1], colonoscopy[2-4], hepato-biliary-pancreatic surgery[5, 6] and inpatient management of chronic diseases such as congestive heart failure.[7] The study of volume-outcome relationships has led to minimum volume standards in coronary artery bypass graft surgery.[8, 9] Intuitively, all high-risk procedures should require a minimum volume standard; however, there are considerable knowledge gaps in certain fields, including endoscopic retrograde cholangiopancreatography (ERCP).

ERCP is one of the most technically complex and high risk endoscopic procedures, the volume-outcome relationship for which is incompletely understood.[10-16] Despite a paucity of data on the topic, experts generally agree that lower volume (endoscopist and facility) is associated with higher failure rates.[17] Therefore, quantifying the volume-outcome relationship for potential targeting of system redesign is increasingly important as U.S. health care policy transitions to a value-based reimbursement or pay-for-performance system.[18] We sought to quantify the relationship between endoscopist volume and failure rates using a regional health information exchange. Our primary aim is to compare failure rates between providers of varying ERCP volume while adjusting for potential confounders. Secondary aims include a presentation of other quality measures, including the rate of diagnostic-only ERCP, post-procedure hospitalization, and 30-day mortality.

Methods

Study Design and Population

We conducted a retrospective cohort study of ERCP procedures identified using insurance claims data derived from the Indiana Network for Patient Care (INPC), a nationally recognized regional health information exchange.[19] Claims data are not restricted to INPC hospitals and include public (Indiana Medicaid) and commercial insurers. Based on membership in Indiana Medicaid and commercial insurance providers included in this cohort, we estimate that 2.28 million Indiana residents (35% of the state population) are represented. Besides insurance claims data, the INPC includes electronic health records data for many facilities ranging from large academic referral centers to community hospitals. Payer claims for ERCP procedures between January, 2001 and December, 2011 were identified using *Current Procedural Terminology*, edition 4 (CPT-4) codes and the

International Classification of Diseases, 9th edition (ICD-9), with each ERCP classified as a distinct event. We validated coding accuracy for the index ERCP and capture of the second ERCP by manual record review of 150 medical records. The study was approved by the Indiana University Office of Research Administration and by the Indiana Office of Medicaid Policy and Planning.

Provider Classification

Endoscopists included gastroenterologists and general surgeons who were classified by their average annual ERCP volume. The list of providers was manually reviewed by three physicians to verify that each provider performed ERCP. To confirm balanced data capture across all providers, we present the average annual number of patient encounters (i.e., all office visits and endoscopic procedures) and number of individual patients having at least one encounter with the provider. Additional data included patient demographics, inpatient or outpatient status at the time of ERCP, and procedure indication (defined using ICD-9 codes).

Outcome Measures

The primary outcome was the post-ERCP failure rate, defined as the composite frequency of the following interventions within 7 days of the index ERCP: having a second (i.e., repeat) ERCP, percutaneous, transhepatic cholangiogram (PTC), or surgery with exploration of the biliary tree (excluding cholecystectomy alone). While one or more of these events does not necessarily equate with a negative outcome, the authors agreed on this composite definition a priori since the frequency of any or all of these events would be negligible if the index ERCP is successful from a diagnostic and therapeutic perspective. The proportion of ERCPs having each of these events is presented individually and as a composite frequency. To avoid double-counting, an index ERCP was defined as an ERCP with no other ERCPs performed within the preceding 30 days. Other than failure rates, additional quality measures included the frequency of purely diagnostic ERCP, defined as an ERCP solely associated with CPT-4 code 43260; post-procedure hospitalization; and all-cause 30-day mortality. We present the frequency and outcomes specific to diagnostic-only ERCP since improvements in less invasive pancreatobiliary imaging such as endoscopic ultrasound and magnetic resonance cholangiopancreatography have nearly eliminated the need for diagnostic-only ERCP. For example, these less invasive imaging tests are preferred to ERCP for patients with a low- or moderate suspicion for choledocholithiasis; if a common bile duct stone is not visualized, ERCP and its inherent risks can be avoided.[20]

Procedure classification

Each ERCP may have more than one associated indication (e.g., gallstone disease + cholangitis). Since an ERCP performed for bile duct diagnosis and therapy is generally considered less complex and has a lower risk profile, indications are grouped into biliary and "other" categories when possible.[21] "Biliary indications" were restricted to ICD-9 codes specific to bile duct pathology: gallstones, bile duct injury or obstruction, cholangitis, and pancreatobiliary malignancy. All other indications, such as acute and chronic pancreatitis, were classified as "other" since the procedure may have been performed for bile duct pathology, pancreatic duct pathology, or both. A table of procedure indications with associated ICD-9 codes is available (see Table, Supplemental Digital Content 1, which lists all relevant ICD-9 codes).

Statistical Analysis

The nature of the volume-outcome association was investigated by fitting a smooth curve on the empirical Bayes estimates of the failure rate using the nonparametric local regression (LOESS) smoothing technique, where the smoothing parameter was selected by the corrected Akaike Information Criterion.[22] The empirical Bayes estimates of the failure rates were estimated based on the logistic regression model with provider-specific random effects where provider volume was modeled as a continuous variable. Providers were then classified by their INPC annual ERCP volume into two relative volume categories, based on the smooth curve that describes the association between provider volume and failure outcome. We present patient, provider and ERCP characteristics along with pre-defined outcomes using descriptive statistics. Continuous variables are summarized using mean and standard deviation if the distribution is roughly symmetric and using median and the 25th and 75th percentiles otherwise. Categorical variables are described using frequency and proportion. Comparison of these characteristics between provider groups is performed using t-tests for provider characteristics and using mixed models with provider-specific random effects for patient and ERCP characteristics/outcomes. Mixed modeling accounts for intra-provider correlation and stabilizes the estimates of provider effects, especially for providers with a small number of ERCPs.

To account for differences in patient characteristics (age, sex, race, inpatient status at the time of ERCP, co-morbidity index), provider characteristics (years since primary board certification and annual volume), and ERCP indication that might have contributed to the observed differences in failure rates across provider groups, these factors were included as covariates in a multivariable model for predicting failure. The multivariable logistic regression model with provider-specific random effects was used to determine the association between provider volume and procedure failure while controlling for the effects of patient- and provider-level covariates. Results are presented using odds ratios (OR), the corresponding 95% confidence intervals (CI), and p values. For each volume strata, the OR represents the odds of procedure failure for every additional ERCP performed per year.

Results

Distribution of ERCP services in Indiana

During the 11-year study period we identified 16,968 ERCPs performed by 130 providers. Of these, 15,514 met criteria for an “index ERCP” while 1,454 (8.6%) had a prior ERCP performed within 30 days. We validated 96% of index ERCPs and 97% of second ERCPs performed within 7 days of the first ERCP. Based on the piecewise linear relationship of relative annual volumes and associated failure rates revealed by the LOESS smooth curve, a cut-off of 25 ERCPs/year was used to dichotomize providers into low (<25/year) and high (≥25/year) (figure 1A-B). There was no difference in the annual number of distinct patient encounters (combination of ERCP and non-ERCP procedures and clinic visits), a surrogate marker of provider-specific clinical activity, captured in the INPC between low and high volume groups (p=0.762), suggesting balanced data capture across these groups (table 1). The majority of providers in both volume groups were gastroenterologists and board-certified. Provider experience, defined as number of years since primary board certification as of 2011, was similar in both groups. The geographic distribution of these providers is illustrated (see Figure, Supplemental Digital Content 2, a map of Indiana illustrating the distribution of ERCP providers across the state.

Patient and Procedure Characteristics

Patients undergoing ERCP by a low volume provider were more likely to be male (p<0.001) but similar in terms of age, inpatient status, and baseline comorbidity as defined using the Charlson co-morbidity index[23] (table 1). While inpatient status at the time of ERCP was similar, patients undergoing ERCP by a low volume provider were more likely to have been admitted one day prior to the procedure (p=0.001) (table 1).

Low volume providers more frequently performed ERCPs for a biliary indication (70.2%, vs. 60.3% among high volume providers), but this did not reach statistical significance (p=0.224) (table 2). However, the proportion of ERCPs performed for acute cholangitis (which typically requires ERCP within 24 hours of presentation) was small in both groups (low – 2.0%, high – 3.0%; p=0.082). High volume providers performed ERCP for bile duct injury, another biliary indication often requiring intervention within 24 hours, more often (0.4%) than the low volume group (0.2%, p=0.216). ERCPs having other indications - those not definitively specific to the biliary tree - were more frequently performed by high volume providers (72.5% vs. 59.7% among low volume; p=0.015). Chronic pancreatitis (p<0.001), sphincter of Oddi dysfunction (p=0.006), and other pancreatic pathology (p<0.001) were substantially more common indications for ERCPs performed by high volume physicians.

Specific indications for ERCP were not associated with greater odds of having a failed procedure among high volume physicians (table 2); in this subgroup, increasing physician volume led to a clinically small but statistically significantly higher chance of failure for ERCPs performed for bile duct obstruction (OR 1.007, p=0.040), while no other indication was sensitive to changes in physician volume. However, among low volume physicians, increasing annual volume was associated with significantly lower odds of a failed ERCP when performed for gallstone disease (OR 0.967, p=0.004), other bile duct/liver pathology (0.963, p=0.002), bile duct obstruction (OR 0.966, p=0.014), abdominal symptoms (0.956, p<0.001), or abnormal laboratory parameters (0.956, p=0.001).

Failure rates and other quality measures

The post-ERCP failure rate, defined as a patient requiring a second ERCP, PTC or surgical exploration of the bile duct ≤ 7 days after the index ERCP, was significantly higher among low volume (9.5%) compared to high volume (5.7%, p<0.001) providers (table 3). The composite rate of early repeat ERCP, PTC, and surgery with bile duct exploration (figure 1A-B) illustrates the piecewise linear volume-outcome relationship. Similarly, the incidence of early (≤ 7 days) repeat ERCP (a subgroup of the definition of procedure failure) was inversely proportional to provider annual volume: low – 4.1%, high – 2.3% (p=0.013). The frequency of diagnostic-only ERCP was significantly higher among low volume (15.8%) compared to high volume providers (7.8%, p=0.016).

Among patients who underwent the procedure as an outpatient, the incidence of immediate hospitalization (defined as hospitalization within 24 hours of the procedure) was significantly higher among low (28.3%) as compared to high volume providers (14.8%, p=0.002). Among patients hospitalized immediately after the ERCP, the length of hospitalization was similar between groups (p=0.899). Similarly, the rates of hospitalization within 30 days of the index ERCP were significantly lower among the highest volume physicians (p=0.023). However, average length of stay was significantly longer for patients admitted within 30 days of ERCP by a high volume provider (p=0.001), suggesting higher baseline morbidity among the subgroup requiring hospitalization. Mortality rates within 30 days of the index ERCP were similar between groups (p=0.455).

Factors associated with greater odds of procedure failure included male sex (OR 1.257, 95% CI: 1.097, 1.440), inpatient status at the time of the procedure (OR 1.521 [1.315, 1.759]), and having a Charlson score ≥ 1 (OR 1.398 [1.058, 1.848]). ERCP indications associated with greater odds of failure included gallstone disease (p=0.028) and bile duct obstruction (p<0.001), whereas acute pancreatitis (p=0.008) and chronic pancreatitis (p=0.001) were associated with a lower failure rate. Physician experience, defined as years since initial board certification, was not significantly associated with a lower rate of failure (p=0.441). Among low volume physicians, the adjusted odds of a failed procedure were significantly lower with each unit increase in procedure volume (OR 0.967 [0.950, 0.984], p<0.001).

Stated alternatively, for each additional ERCP performed per year by a low volume provider, the odds of failure decreased by 3.3% (95% CI 1.6-5.0%). Differences in procedure volume among high volume providers did not impact odds of failure (p=0.997).

Conclusions

Over the past two decades, there has been a slow increase in the number of patients referred for ERCP to an academic medical center after having a failed procedure at a lower-volume facility.[24] Therefore, we emphasize failure rates, frequency of diagnostic-only ERCP, and need for hospitalization as key quality measures in an effort to define the endoscopist volume-outcome relationship. Previous studies evaluating this relationship are conflicting.[11-14, 25] These studies are limited by voluntary reporting of outcomes,[14] emphasis on procedural complications that are highly dependent on the pre-procedure indication, and sampling frames with small numbers of physicians or facilities. Varadarajulu and colleagues previously reported higher failure rates and longer hospitalizations for ERCPs performed at lower volume facilities but did not consider individual provider volume.[12] In their study, failure was defined as the need for PTC or bile duct exploration but did not include the rate of early repeat ERCP, an important measure of incomplete therapy. In addition, Varadarajulu, et al. restricted their analysis to select ICD-9 codes (cholelithiasis, other disorders of the biliary tract, and diseases of the pancreas) and did not include purely diagnostic ERCPs. In our cohort, failure rates were higher among purely diagnostic ERCPs, presumably because the majority of diagnostic ERCPs represent cases in which one or both ducts can be opacified but the physician is unable to proceed with necessary therapy (e.g., sphincterotomy, stone removal, stent placement) for technical reasons. The alternative explanation is that the procedure was performed for diagnosis alone. In either case, a diagnostic-only ERCP represents lower quality care. Therefore, our observation that rates of purely diagnostic ERCP are higher among low volume providers is another important measure of quality, since the need for diagnostic ERCP has been obviated by alternative, less invasive imaging modalities. The study by Varadarajulu, et al. observed a facility volume-outcome relationship that is congruent with our results at the provider level. It is unlikely that correcting for facility volume in our analysis would eliminate the observed provider outcome relationship since provider and facility volume are often closely linked: the majority of low volume providers perform ERCP typically at low volume facilities.

There are several compelling reasons that the “value” (i.e., outcome for each dollar spent[18]) of ERCP services in the U.S. is low. First, its utilization has plateaued over the last decade since it has been largely replaced as a purely diagnostic tool by less invasive modalities such as endoscopic ultrasound and magnetic resonance cholangiopancreatography.[26, 27] As a result, there are comparatively fewer procedures to disperse among ERCP providers who need to develop and maintain their technical expertise. Second, the majority (95%) of ERCPs performed in the United States are completed in lower volume (<200 per year) facilities.[12] Low volume units are less efficient than high volume units, which can achieve a lower average cost. Third, ERCP is performed less frequently than other endoscopic procedures, and is frequently required in an urgent setting (e.g., symptomatic choledocholithiasis or obstructive jaundice). Large (> 5 providers) specialty practices are reluctant to centralize ERCP services in part due to the implications for night and weekend coverage.[28] However, there is almost invariably an opportunity to delay the procedure for a limited time since indications for emergent (< 24 hours after presentation) ERCP such as cholangitis (2.6% in our cohort) and some cases of bile duct injury (0.3% in our cohort) are uncommon. Since the current U.S. healthcare delivery system has traditionally incentivized quantity over quality, hospitals presumably choose to offer ERCP in order to benefit from “downstream revenue:” utilization of cross-sectional imaging, surgery, and hospitalization. However, in the era of healthcare reform and patient-reported outcomes, a more conscious effort to balance cost and quality will continue to evolve. Fourth, ERCP is one of the most technically difficult endoscopic procedures with a high risk profile, and most gastroenterologists who perform it have limited training due to the aforementioned factors.[28] Finally, as shown in our cohort, the majority (116 of 130, 89%) of physicians performing ERCP do very little of it.

Particular strengths of our study are the inclusion of both commercial and public payers; given the indications for ERCP such as choledocholithiasis, this population is more generalizable than a Medicare population of individuals > age 65. Another advantage is the large sample size across the diverse demography of Indiana, derived from a nationally recognized health information exchange. Our inclusion of all reported indications for ERCP improves the generalizability of the results and underscores that, despite higher failure and post-procedure hospitalization rates, low volume providers are more likely to perform ERCP for indications that typically have lower complexity and complication rates. Therefore, it is possible that our observed differences are biased against high volume providers. However, use of ICD-9 codes to define indications does not necessarily distinguish high and low complexity procedures: a 4mm gallstone that can be easily removed following biliary sphincterotomy would have the same code as a 20mm stone that requires mechanical or electrohydraulic lithotripsy. Finally, our definition of failure is specific but has lower sensitivity since a substantial group of patients may have required a repeat ERCP, PTC or bile duct exploration after 7 days due to an unsuccessful index procedure (e.g., painless jaundice, choledocholithiasis without acute symptoms). It is very unlikely that we have overestimated the failure rate since a second bile duct intervention in ≤ 7 days after ERCP is almost never clinically indicated if the index procedure is successful; specific cases such as Mirizzi syndrome may require surgical exploration after a successful ERCP but this is usually deferred several weeks to allow normalization of cholestatic liver tests. Furthermore, changing the window for failure to 15 or 30 days did not significantly change the results; a significant relationship between low volume and higher failure rates persisted (data not shown).

There are important limitations to consider. First, we did not quantify the interaction between physician and facility volume since the latter could not be reliably measured. Second, we did not measure other physician-specific factors that may affect outcomes, such as annual volume prior to the onset of this study and quality of ERCP training. Consistent with our analysis, a previous study of 1,335 ERCPs observed lower complication rates from sphincterotomy among higher volume providers, irrespective of experience, defined as years since training.[29] Third, we do not present rates of ERCP-specific complications such as pancreatitis, gastrointestinal bleeding, cholangitis and bowel perforation; these are expected to correlate with procedure indication more than to our definition of failure. However, our observation that hospitalization rates within 24 hours of the procedure are higher among low volume providers can be explained either by a tendency among low volume providers to admit more patients for observation post-procedure or by a higher incidence of post-procedure complications. In either case, higher post-procedure hospitalization rates suggest lower quality of care, excess resource utilization, or both, but require confirmation through more direct study. Finally, while the INPC data is high quality[19] it does not capture all ERCPs performed by each provider. Interestingly, we found similar rates of overall clinical productivity, defined as total number of clinical encounters per provider, between the high and low ERCP volume provider groups, which suggests that we are not misclassifying providers as low volume simply on the basis of missing data. We estimate that the observed INPC volumes imply overall annual volumes of < 117 ERCPs/year for low volume providers. However, INPC-specific volumes should not be considered as “thresholds” for future guidelines/recommendations without more direct study. A low volume cut-off of 117 procedures is consistent with previous studies suggesting lower complication rates among physicians performing > 100 ERCPs annually.[29] This has important implications for systems redesign, since we observed similar outcomes for all providers exceeding this threshold.

In conclusion, despite performing ERCP more frequently for indications having greater technical complexity and risk, higher volume providers have lower failure rates and hospitalization rates compared to lower volume physicians. The magnitude of this volume-outcome relationship highlights the importance of developing benchmark reports for ERCP outcomes that are provider- and facility-specific, as the Gastrointestinal Quality Improvement Consortium (GIQUiC) developed for colonoscopy.[30] The consequences that ERCP-specific outcome reporting might have upon current patient preferences to receive care locally and the potential to widen or to narrow disparities in care require further investigation.[31-33] ERCP outcomes may be similar to pancreatectomy, the mortality from which significantly declined as the surgery was performed in more selective centers.[34] In addition to performance transparency, a recent Institute of Medicine report emphasized the need to utilize systems engineering tools and process improvement methods to provide the best care at lower cost.[35] Efforts to concentrate ERCP services among fewer, higher volume providers may help to achieve this goal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: NIH/NIDDK K23DK095148 (Coté)

Abbreviations

ERCP	endoscopic retrograde cholangiopancreatography
INPC	Indiana Network for Patient Care
CPT	Current Procedural Terminology
ICD	International Classification of Diseases
PTC	percutaneous transhepatic cholangiography
OR	odds ratio
LOESS	local regression smoothing technique

References

1. Ananthakrishnan AN, McGinley EL, Saeian K. Higher hospital volume is associated with lower mortality in acute nonvariceal upper-GI hemorrhage. Gastrointest Endosc. 2009; 70:422–32. [PubMed: 19560760]
2. Bernstein C, Thorn M, Monsees K, et al. A prospective study of factors that determine cecal intubation time at colonoscopy. Gastrointest Endosc. 2005; 61:72–5. [PubMed: 15672059]
3. Radaelli F, Meucci G, Sgroi G, et al. Technical performance of colonoscopy: the key role of sedation/analgesia and other quality indicators. Am J Gastroenterol. 2008; 103:1122–30. [PubMed: 18445096]
4. Singh H, Penfold RB, DeCoster C, et al. Colonoscopy and its complications across a Canadian regional health authority. Gastrointest Endosc. 2009; 69:665–71. [PubMed: 19251007]
5. Bilimoria KY, Talamonti MS, Sener SF, et al. Effect of hospital volume on margin status after pancreaticoduodenectomy for cancer. J Am Coll Surg. 2008; 207:510–9. [PubMed: 18926452]

6. Nathan H, Cameron JL, Choti MA, et al. The volume-outcomes effect in hepato-pancreato-biliary surgery: hospital versus surgeon contributions and specificity of the relationship. J Am Coll Surg. 2009; 208:528–38. [PubMed: 19476786]
7. Joynt KE, Orav EJ, Jha AK. The association between hospital volume and processes, outcomes, and costs of care for congestive heart failure. Ann Intern Med. 2011; 154:94–102. [PubMed: 21242366]
8. Peterson ED, Coombs LP, DeLong ER, et al. Procedural volume as a marker of quality for CABG surgery. JAMA. 2004; 291:195–201. [PubMed: 14722145]
9. Shahian DM. Improving cardiac surgery quality--volume, outcome, process? JAMA. 2004; 291:246–8. [PubMed: 14722153]
10. 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 cholangio-pancreatograph practice. Gut. 2007; 56:821–9. [PubMed: 17145737]
11. 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]
12. Varadarajulu S, Kilgore ML, Wilcox CM, et al. Relationship among hospital ERCP volume, length of stay, and technical outcomes. Gastrointest Endosc. 2006; 64:338–47. [PubMed: 16923479]
13. 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]
14. 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
15. Freeman ML, DiSario JA, Nelson DB, et al. Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. Gastrointest Endosc. 2001; 54:425–34. [PubMed: 11577302]
16. Loperfido S, Angelini G, Benedetti G, et al. Major early complications from diagnostic and therapeutic ERCP: a prospective multicenter study. Gastrointest Endosc. 1998; 48:1–10. [PubMed: 9684657]
17. Dumonceau JM, Andriulli A, Deviere J, et al. European Society of Gastrointestinal Endoscopy (ESGE) Guideline: prophylaxis of post-ERCP pancreatitis. Endoscopy. 2010; 42:503–15. [PubMed: 20506068]
18. Porter ME. A strategy for health care reform--toward a value-based system. N Engl J Med. 2009; 361:109–12. [PubMed: 19494209]
19. McDonald CJ, Overhage JM, Barnes M, et al. The Indiana network for patient care: a working local health information infrastructure. An example of a working infrastructure collaboration that links data from five health systems and hundreds of millions of entries. Health Aff (Millwood). 2005; 24:1214–20. [PubMed: 16162565]
20. Scheiman JM, Carlos RC, Barnett JL, et al. Can endoscopic ultrasound or magnetic resonance cholangiopancreatography replace ERCP in patients with suspected biliary disease? A prospective trial and cost analysis. Am J Gastroenterol. 2001; 96:2900–4. [PubMed: 11693324]
21. Baron TH, Petersen BT, Mergener K, et al. Quality indicators for endoscopic retrograde cholangiopancreatography. Gastrointest Endosc. 2006; 63:S29–34. [PubMed: 16564909]
22. Hurvich CM, Simonoff JS, Tsai CL. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. Journal of the Royal Statistical Society Series B-Statistical Methodology. 1998; 60:271–293.
23. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987; 40:373–83. [PubMed: 3558716]
24. Cote GA, Singh S, Bucksot LG, et al. Association between volume of endoscopic retrograde cholangiopancreatography at an academic medical center and use of pancreatobiliary therapy. Clin Gastroenterol Hepatol. 2012; 10:920–4. [PubMed: 22387254]
25. 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]
26. Mazen Jamal M, Yoon EJ, Saadi A, et al. Trends in the utilization of endoscopic retrograde cholangiopancreatography (ERCP) in the United States. Am J Gastroenterol. 2007; 102:966–75. [PubMed: 17319932]

27. Peery AF, Dellon ES, Lund J, et al. Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterology. 2012; 143:1179–1187. e3. [PubMed: 22885331]
28. Cote GA, Keswani RN, Jackson T, et al. Individual and practice differences among physicians who perform ERCP at varying frequency: a national survey. Gastrointest Endosc. 2011; 74:65–73. e12. [PubMed: 21492851]
29. Rabenstein T, Schneider HT, Nicklas M, et al. Impact of skill and experience of the endoscopist on the outcome of endoscopic sphincterotomy techniques. Gastrointest Endosc. 1999; 50:628–36. [PubMed: 10536317]
30. GI Quality Improvement Consortium. Volume 2012
31. Birkmeyer JD, Siewers AE, Marth NJ, et al. Regionalization of high-risk surgery and implications for patient travel times. JAMA. 2003; 290:2703–8. [PubMed: 14645312]
32. Liu JH, Zingmond DS, McGory ML, et al. Disparities in the utilization of high-volume hospitals for complex surgery. JAMA. 2006; 296:1973–80. [PubMed: 17062860]
33. Schwartz LM, Woloshin S, Birkmeyer JD. How do elderly patients decide where to go for major surgery? Telephone interview survey. BMJ. 2005; 331:821. [PubMed: 16192286]
34. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. N Engl J Med. 2011; 364:2128–37. [PubMed: 21631325]
35. Medicine Io. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Institute of Medicine of the National Academies; 2012.

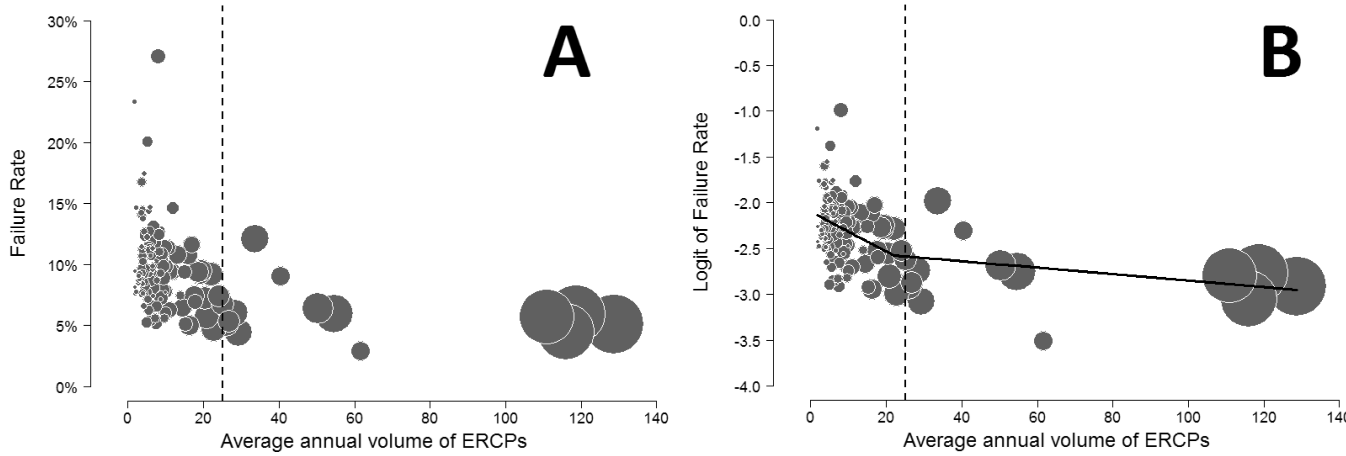

Figure 1A-B. Composite Failure Rate of ERCP by Provider Volume

Scatterplot depicting composite rates of repeat ERCP, PTC and surgery with bile duct exploration performed within 7 days of the index ERCP, with dotted vertical line delineating volume groups (1A). Frequencies are given using Bayes estimates. The failure rate is inversely proportional to provider annual volume, confirming a volume-outcome relationship. The size of each bubble is proportional to the number of ERCPs performed. The x axis reflects annual ERCP volume based on INPC data. Since the logistic regression model assumes the logit of the outcome as a linear function of the predictors, we show the logit of the failure rate estimate against physician volume (1B). The solid line is the nonparametric smoothing curve that best describes the relationship between physician volume and the logit of the failure rate.

Table 1

Patient and Provider Characteristics

Variable	Annual volume[†]			
	Low	High	Total	P value
Provider characteristics				
	n=116 Providers	**n=14 Providers**	**n=130 Providers**	
Annual ERCP volume (# ERCPs/year), mean (SD)	7.9 (5.1)	60.7 (39.8)	13.6 (21.3)	
Total number of distinct patient encounters/year, mean (SD)	817.7 (568.4)	866.0 (514.8)	822.9 (561.2)	0.762
Total number of patients/year, mean (SD)	415.5 (201.5)	443.2 (222.9)	418.5 (203.2)	0.632
Physician specialty, n (%) Gastroenterologist	106 (91.4%)	14 (100%)	120 (92.3%)	0.599
Board Certification	106 (94.6%)	10 (83.3%)	116 (93.6%)	0.173
Missing, n	4	2	6	
Years of board certification, as of 2011, mean (sd)	17.1 (9.0)	18.2 (8.5)	17.2 (9.0)	0.701
Missing, n	8	1	9	
Patient characteristics				
	n=7,484 ERCPs	**n=8,030 ERCPs**	**n=15,514 ERCPs**	
Inpatient status at the time of ERCP, n (%)	3,138 (41.9%)	2,631 (32.8%)	5769 (37.2%)	0.198
• Admitted one day before ERCP	1,122 (15.0%)	655 (8.2%)	1777 (11.5%)	0.001
• Admitted 2+ day before ERCP	2,016 (26.9%)	1,976 (24.6%)	3992 (25.7%)	0.879
Patient age, mean (sd)	53.2 (20.3)	50.0 (18.4)	51.5 (19.4)	0.547
Missing, n	638	766	1404	
Male sex, n (%)	1,994 (29.1%)	2,678 (36.9%)	4,672 (33.1%)	<0.001
Missing, n	639	766	1405	
Ethnic group, n (%)				0.240
White	4,618 (81.5%)	6,016 (85.1%)	10,634 (83.5%)	
African American	644 (11.4%)	727 (10.3%)	1,371 (10.8%)	
Other	406 (7.2%)	326 (4.6%)	732 (5.7%)	
Missing, n	1,816	961	2777	
Charlson Co-morbidity Index				0.169
0	7,162 (95.7%)	7,683 (95.7%)	14,845 (95.7%)	
1+	322 (4.3%)	347 (4.3%)	669 (4.3%)	

[†]Annual volume was defined by the average annual number of ERCPs performed by an individual provider in the INPC. Low annual volume: ≤ 25 ERCPs/year; high annual volume: >25 ERCPs/year.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Effect of physician volume on 7-day failure outcome, stratified by indication for ERCP

Procedure Indication	Low (n=116 providers with 7,484 ERCPs)			High (n=14 providers with 8,030 ERCPs)			Total (n=130 providers with 15,514 ERCPs)
	N (%)	Odds Ratio	P value	N (%)	Odds Ratio	P value	
Biliary Indications	5,254 (70.2%)			4,842 (60.3%)			10,096 (66.4%)
• Gallstone disease	3,211 (42.9%)	0.967	0.004	1,926 (24.0%)	0.998	0.623	5,137 (33.1%)
• Other bile duct/liver pathology	1,312 (17.5%)	0.963	0.002	2,158 (26.9%)	0.999	0.829	3,470 (22.4%)
• Bile duct obstruction	1,262 (16.9%)	0.966	0.014	1,785 (22.2%)	1.007	0.040	3,047 (19.6%)
• Cholangitis	152 (2.0%)			244 (3.0%)			396 (2.6%)
• Bile duct injury	12 (0.2%)			34 (0.4%)			46 (0.3%)
• Malignancy, Pancreas	118 (1.6%)			148 (1.6%)			266 (1.7%)
• Malignancy, Bile Duct/ Liver	77 (1.0%)			92 (1.2%)			169 (1.1%)
Other Indications	4,468 (59.7%)			5,819 (72.5%)			10,287 (67.7%)
• Abdominal symptoms	2,707 (36.2%)	0.956	<0.001	3,069 (38.2%)	0.996	0.182	5,776 (37.2%)
• Abnormal laboratory parameters	1,474 (19.7%)	0.956	0.001	1,401 (17.5%)	1.006	0.077	2,875 (18.5%)
• Acute pancreatitis	735 (9.8%)	0.989	0.610	808 (10.1%)	0.992	0.118	1,543 (10.0%)
• Chronic pancreatitis	356 (4.8%)	0.972	0.391	1,229 (15.3%)	1.000	0.979	1,585 (10.2%)
• Malignancy NOS	75 (1.0%)			174 (2.2%)			249 (1.6%)
• Other pancreatic pathology	240 (3.2%)			1,232 (15.3%)			1,472 (9.5%)
• Sphincter of Oddi dysfunction	173 (2.3%)			1,158 (14.2%)			1,331 (8.6%)
Other (none of the above)	178 (2.4%)			136 (1.7%)			314 (2.0%)

• Each procedure may have more than one indication, so numbers do not add up to 100%.

• Odds ratio represents the odds of having a failed procedure for an increase in annual ERCP volume of n=1 among low and high volume physicians, respectively. For example, among low volume physicians performing an ERCP for gallstone disease, the odds of having a failed procedure are 3.3% lower for each increase of one ERCP per year.

• With the exception of bile duct obstruction, the odds of having a failed procedure did not significantly change among high volume providers per unit increase in ERCP volume.

Table 3

ERCP Quality Measures

Outcome	Low (n=116 providers with 7,484 ERCPs)	High (n=14 providers with 8,030 ERCPs)	Total (n=130 providers with 15,514 ERCPs)	P value*
7-day Failure rate – Total	707 (9.5%)	456 (5.7%)	1163 (7.5%)	<0.001
• Early repeat ERCP	306 (4.1%)	188 (2.3%)	494 (3.2%)	0.013
• PTC or surgery	442 (5.9%)	283 (3.5%)	725 (4.7%)	0.001
Diagnostic-only ERCP	1,182 (15.8%)	627 (7.8%)	1809 (11.7%)	0.016
• 7-day repeat ERCP rate	118 (10.0%)	40 (6.4%)	158 (8.7%)	0.091
• 7-day rate of PTC or surgery	108 (9.1%)	40 (6.4%)	148 (8.2%)	0.096
• 7-day composite failure rate	209 (17.7%)	77 (12.3%)	286 (15.8%)	0.049
Immediate hospitalization after ERCP†¥	1,231 (28.3%)	801 (14.8%)	2032 (20.9%)	0.002
• Length of stay, days (median, 25th and 75th percentile)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	3.0 (2.0, 5.0)	0.899‡
Hospitalization in 30 days after ERCP†	1,811 (41.7%)	1,641 (30.4%)	3452 (35.4%)	0.023
• Length of stay, days (median, 25th and 75th percentile)	4.0 (2.0, 7.0)	4.0 (2.0, 10.0)	4.0 (2.0, 7.0)	0.001‡
30-day mortality	140 (1.9%)	154 (1.9%)	294 (1.9%)	0.455

*P-values are obtained by adjusting the clustering effect of providers using a logistic regression model with random effects.

†Hospitalization data are limited to patients who were not inpatients at the time of ERCP (n=9,745).

¥Immediate hospitalization is defined as a hospitalization that occurred within 24 hours of the index ERCP.

‡Because length of stay is highly skewed to the right, comparisons are based on its logarithmic transform.

Table 4

Multivariate analysis: Risk of ERCP failure

	Odds Ratio	95% Confidence Interval		P-value
Patient characteristics				
Age	0.999	0.996	1.003	0.746
Gender, male vs. female	1.257	1.097	1.440	0.001
Race: Black vs. White	1.003	0.802	1.255	0.978
Race: Other vs. White	0.961	0.721	1.282	0.786
Race: Missing vs. White	0.897	0.712	1.131	0.355
Inpatient at ERCP	1.521	1.315	1.759	<0.001
Inpatient at ERCP: admitted one day before ERCP	0.892	0.730	1.090	0.260
Charlson Comorbidity Index, 1+ vs. 0	1.398	1.058	1.848	0.019
Gallstone Disease	1.182	1.018	1.371	0.028
Other bile duct/liver pathology	1.129	0.963	1.323	0.134
Bile duct obstruction	1.486	1.253	1.761	<0.001
Abdominal symptoms	1.040	0.901	1.201	0.589
Abnormal laboratory parameters	1.101	0.933	1.299	0.254
Acute pancreatitis	0.698	0.536	0.909	0.008
Chronic pancreatitis	0.566	0.414	0.773	0.001
Provider characteristics				
Physician specialty – Gastroenterologist *	-	-	-	-
Board certification *	--	-	-	-
Years since primary board certification	0.994	0.980	1.008	0.411
Physician volume, when < 25	0.967	0.950	0.984	<0.001
Physician volume, when 25+	1.000	0.994	1.006	0.997

* Physician specialty and board certification are not included in the analysis because 92% of physicians are Gastroenterologists, and 90% are board certified