



Published in final edited form as:

*Gastrointest Endosc.* 2018 January ; 87(1): 164–173.e2. doi:10.1016/j.gie.2017.04.030.

## Provider Specific Quality Measurement for Endoscopic Retrograde Cholangiopancreatography Utilizing Natural Language Processing

Timothy D. Imler<sup>1,2,3</sup>, Stuart Sherman<sup>1,2</sup>, Thomas F. Imperiale<sup>1,2,4,5</sup>, Huiping Xu<sup>6</sup>, Fangqian Ouyang<sup>6</sup>, Christopher Beesley<sup>3</sup>, Charity Hilton<sup>3</sup>, and Gregory A. Côté<sup>1,7</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology; Indiana University School of Medicine. Indianapolis, IN

<sup>2</sup>Department of Medicine; Indiana University School of Medicine. Indianapolis, IN

<sup>3</sup>Department of Biomedical Informatics; Regenstrief Institute, LLC. Indianapolis, IN

<sup>4</sup>Center of Innovation; Health Services Research and Development; Richard L. Roudebush VA Medical Center. Indianapolis, IN

<sup>5</sup>Health Services Research, Regenstrief Institute, LLC. Indianapolis, IN

<sup>6</sup>Department of Biostatistics; Indiana University School of Medicine. Indianapolis, IN

<sup>7</sup>Division of Gastroenterology and Hepatology; Medical University of South Carolina. Charleston, SC

### Abstract

**Background**—Natural language processing (NLP), an information retrieval technique from text documents, accurately identifies quality measures for other endoscopic procedures. There are no systematic methods by which to track adherence to quality measures for endoscopic retrograde cholangiopancreatography (ERCP), the highest risk endoscopic procedure widely utilized in practice.

---

Correspondence: Timothy D. Imler, M.D., M.S., Assistant Professor of Medicine; Division of Gastroenterology and Hepatology Research Scientist; Regenstrief Institute 702 Rotary Circle; Suite 225; Indianapolis, Indiana 46202, Office: (317) 274-9137; Fax: (317) 278-6870; timler@iu.edu.

#### Contributions

Study design (Imler/Cote/Imperiale/Sherman); Data collection (Imler/Hilton/Beesley); Data analysis (Imler/Cote/Xu/Ouyang/Sherman); Statistical analysis (Imler/Xu/Ouyang/Imperiale); Manuscript drafting (Imler); Critical editing (All listed authors)

#### Conflicts of Interest

The authors disclose that Dr. Imler has filed for provisional patent (IURTC-14098-01-US-E) for similar work (colonoscopy quality) under the name Tracking Real-time Assessment of Quality Monitoring in Endoscopy (TRAQ-ME) through the Indiana University Research and Technology Corporation (IURTC).

**Duplicate/Previous Publications:** None

#### Ethics

This study was approved by the Institutional Review Board at Indiana University. Drs. Imler has filed for provisional patent (IURTC-14098-01-US-E) for this work under the name Tracking Real-time Assessment of Quality Monitoring in Endoscopy (TRAQ-ME) through the Indiana University Research and Technology Corporation (IURTC). No other conflicts of interest are claimed by the remaining authors.

**Aim**—Demonstrate the feasibility of NLP to track adherence to ERCP quality measures across individual providers.

**Methods**—Six providers at a single institution had their ERCPs identified from 2006–2014. Quality measures were defined using society guidelines and extracted using a combination of NLP and data mining (e.g. ICD-9 CM codes). Validation for each quality measure was performed by manual record review. Individualized quality measures were compared across the six providers in analyses adjusted and unadjusted for patient age, sex, and race. Quality measures were grouped into pre-procedure (5), intra-procedure (6), and post-procedure (2). NLP was evaluated utilizing measures of precision and accuracy.

**Results**—23,674 ERCPs were included in the analysis (average patient age of 52.9±17.8) with 14,113 (59.6%) women.

Among the thirteen quality measures, precision of NLP ranged from 84–100% with intra-procedure measures having lower precision (84% for precut sphincterotomy). Accuracy of NLP ranged from 90–100% with intra-procedure measures having lower accuracy (90% for pancreatic stent placement). One provider did not meet the “appropriate indication” quality measure within their individualized 95% confidence interval (78.4 (77.1–79.9)). Documentation of adverse events showed the greatest variation (8.8–92.6%).

**Conclusion**—Use of NLP and data mining allow for individualized tracking of ERCP providers for quality metrics without requiring manual medical record review. Incorporation of these tools across multiple centers may demonstrate the ability to track ERCP quality measures on a regional or national level.

### Keywords

Endoscopic retrograde cholangiopancreatography; quality measurement; natural language processing

## Background

Quality measurement of endoscopy is becoming the standard of care in the United States<sup>1–5</sup> and may influence choice in provider, outcomes, and reimbursement<sup>5–7</sup>. Endoscopic retrograde cholangiopancreatography (ERCP), the highest risk endoscopic procedure in widespread practice<sup>8</sup> has not been extensively studied for individual, endoscopic-based quality measures<sup>9</sup>. Historically, ERCP quality has been focused on provider or facility volume, with higher volumes associated with higher quality as defined by success and complication rates after adjusting for procedure indication<sup>10, 11</sup>. In 2006, the American Society for Gastrointestinal Endoscopy (ASGE) and the American College of Gastroenterology (ACG) Task Force on Quality Endoscopy provided the first quality indicators for ERCP, many based on expert consensus<sup>9</sup>, and these were subsequently updated in 2014<sup>1</sup>.

Similar to colonoscopy, there are challenges in obtaining ERCP-based quality measures due to the time intensive nature of manual medical record review. Given this challenge in colonoscopy, several studies have demonstrated the feasibility of using natural language

processing (NLP) to extract these measures from text documents in the medical record, with > 90% accuracy for colonoscopy-specific measures<sup>12–16</sup>. We hypothesized that NLP could be used to track ERCP quality measures accurately and efficiently. If successful, NLP could be employed across health systems to monitor ERCP quality and provide feedback to providers, administrators, and payers in an effort to show adherence to national benchmarks and, if needed, refinement following quality improvement interventions<sup>17</sup>.

Given the need to systematic measure adherence to quality benchmarks specific to ERCP, the primary aim of this study is to measure the precision and accuracy of NLP in automatically measuring adherence to ERCP-specific quality measures. The secondary aim is to quantify the variation of adherence to quality benchmarks among providers at a single institution.<sup>17</sup>

## Methods

After Institutional Review Board (IRB) and Regenstrief Institute data management committee approval, we identified ERCP procedure reports and related clinical data from 1/1/2006 through 7/25/2014.

### Data Source

The Indiana Network for Patient Care (INPC)<sup>18</sup> is a large regional health information exchange that obtains data from 25,000 physicians, 94 hospitals, 110 clinics and surgery centers and other healthcare organizations as well as payer data<sup>19, 20</sup>. The database houses more than 4 billion pieces of clinical data with over 160 million text reports and has stored data for over 40 years. All ERCP procedure reports as well as ERCP-related radiology reports are stored within the INPC and are grouped as the “ERCP” document type and were created utilizing a single endoscopy software (Provation® MD; Wolters Kluwer). Clinical and payer data sources facilitate pairing ERCP reports with procedure indications (International Classification of Diseases, Ninth Revision, Clinical Modification, ICD9-CM) and specific maneuvers (Current Procedural Terminology, CPT) codes).

### Natural Language Processing System

The Regenstrief Institute has created an Apache Unstructured Information Management Applications (UIMA<sup>TM</sup>)<sup>21</sup> based NLP system (nDepth) that utilizes open-source applications for NLP processing and is released under the Apache license version 2.0<sup>22</sup>. The NLP system houses more than 90 million text-based electronic documents from multiple institutions throughout Indiana. All documents are indexed within Apache Solr<sup>TM</sup><sup>23</sup> (a method for providing distributed indexing with load-balanced querying) for access via Boolean search (e.g. “AND”, “OR”, “NOT”). In addition to search, more advanced NLP techniques (e.g. negation, regular expressions, and standard terminologies) are also available through the system.

### Quality Measure Selection

Quality measures were identified based on the 2014 ASGE/ACG Quality Indicators for ERCP that were reviewed and endorsed by the American Society of Gastrointestinal

Endoscopy (ASGE), the American College of Gastroenterology (ACG), and the American Gastroenterological Association (AGA)<sup>1</sup>. Additional measures (e.g. pre-cut sphincterotomy utilized for cannulation) were added from internal discussion with content experts. Measures were categorized as: 1) pre-procedure, 2) intra-procedure, and 3) post-procedure. Table 1 lists the quality indicators and the method by which they were identified. Use of rectal indomethacin was not selected as a quality measure since this would have required additional programming to interface with pharmacy databases.

### Validation of Quality Measures by NLP

We randomly separated the cohort 1:1 into training and testing sets. The training set of ERCP reports was used to create NLP algorithms in an iterative fashion (Appendix Figure 1). The test documents (which were not reviewed during the algorithm development phase) were then used to evaluate the ability of the system to accurately determine the expected finding. Developing independent training and test sets was done to avoid 'over-fitting' the algorithm which could reduce external validity.

NLP was utilized to identify ERCP endoscopist and all listed quality measures in Table 1. For each quality measure, we randomly selected 50 documents from the training set for which the quality measure was identified by NLP and 50 others for which the quality measure was not identified. A single expert gastroenterologist (TDI) reviewed all documents (different random selection for each quality measure) from the NLP search and assessed true positives (TP, those documents that were appropriately identified by the search) and true negatives (TN, those documents without the presence of the quality measure by NLP). The precision (True positives/50) and accuracy of the NLP search for each quality measure were assessed based on these manually reviewed documents. We estimated that 50 documents were sufficient for procedure identification given the high prevalence of the measures within the dataset as well as our previous experience with NLP validation within endoscopy<sup>12, 25</sup>. A formal power calculation was not performed due to inability to know the percentage of the various outcomes *a priori*.

Statistical analysis was performed to quantify the precision and accuracy of the individual quality measures by NLP. Since the entire document set (n = 63,119 ERCP procedures) was not manually annotated for a gold standard, a true sensitivity (reports in agreement/positive reports by manual review) could not be calculated for this study.

**Precision:** True positives/Test outcome positives over the 50 reviewed documents.

**Accuracy:** True positives + True negatives/Total population of 100 reviewed documents.

### Extraction of non-NLP based Quality Measures (Metrics 1, 6–9, and 12)

Quality measures that did not require text extraction via NLP were extracted/obtained by a data manager for the INPC and are listed in Table 1. These quality measures were extracted using ICD-9-CM codes and linked to individual ERCP reports using the master medical record number and date of procedure.

Metric “1” (appropriateness of indication) was searched according to ICD-9 CM codes based on the multi-society recommendations for appropriate indications<sup>1</sup>. Appropriate codes were identified by rating the indication as highly appropriate or potentially appropriate. These codes were rated if they had more than 100 instances of being used in the first or second position of billing within 7 days of an ERCP procedure. Additional codes were added based on the known appropriate indications despite having less than 100 instances of being utilized. ICD9-CM codes for Metric 1 are listed in the appendix. Metric “12” for perforation related to the procedure was determined based on ICD-9 CM codes including perforation of bile duct (576.3) and perforation of intestine (569.83) within 7 days of ERCP procedure. The metric rate was calculated as the number of perforations/the number of total ERCP by provider and in aggregate. Metric “13” for significant post-sphincterotomy bleeding was determined based on ICD-9 code acute post-hemorrhagic anemia (285.1) within 7 days of procedure. The metric rate was # of bleeding events/# of total ERCP by provider and in aggregate.

Metrics 6–9 required ICD-9 CM identification prior to NLP analysis. These measures were defined using the subgroup of ERCPs performed for choledocholithiasis. Inclusion codes were calculus of bile duct (574.3\*, 574.4\*, 574.5\*) with exclusion codes of spasm of sphincter of Oddi (576.5), acute pancreatitis (577.1), 577.2 (cyst and pseudocyst of pancreas), 577.1 (chronic pancreatitis), 577.8 (other specified disease of pancreas), 751.7 (anomalies of pancreas), and 157.\* (malignant neoplasm of pancreas). We selected choledocholithiasis for these metrics since this indication is one of the most common and requires biliary cannulation, common bile duct stone extraction, and rates of inadvertent pancreatic duct cannulation and injection denote technical proficiency in safely executing the procedure.

### Comparison of Individual Provider Quality Measures

After each provider (n=6) was identified by NLP, individualized quality metrics were extracted. For each provider, patient characteristics including age, gender, and race were summarized using mean and standard deviations for continuous variables and frequency and proportions for categorical variables. For binary quality metrics, unadjusted rates for each provider were calculated using proportions and compared using the Pearson chi-square test. Adjusted rates of binary quality metrics were obtained from a logistic regression model that controlled for patient characteristics including age, gender, and race. For the number of ERCPs performed per year, mean and standard deviation of annual procedural volume were calculated over the study period; providers were compared using the ANOVA F-test. All statistical analyses were performed using SAS 9.4 (SAS Institute, Cary, NC).

### Results

Of 63,119 documents on 15,581 patients that were indexed as an “ERCP” document, 39,440 were excluded as they were radiology reports or other non-procedure reports. This resulted in 23,679 ERCP procedures on 13,299 patients identified by NLP (figure 1). Validation of this methodology showed all reviewed documents to be true ERCP procedure reports written by an endoscopist (e.g. not a radiology report). Of 23,679 ERCP procedures, 5 were missing

patient age, gender, or race and hence were excluded as having incomplete data and not available for adjustment. The remaining 23,674 procedures were utilized as the final study sample to evaluate the quality measures across the six providers.

The mean age of patients at the time of the procedure was  $52.9 \pm 17.8$  with 59.6% female (Appendix Table 1). The majority (75.9%) of patients were Caucasian. Table 2 shows the primary outcome of validation including precision and accuracy of the quality measures. Precision ranged from 84–100% with intra-procedure measures having lower precisions. Accuracy ranged from 90–100% with intra-procedure measures having lower accuracy. We excluded seventeen documents post-hoc because the primary provider listed on the note was a trainee ( $n=9$ ) or rarely performed ERCP ( $n=8$ ).

Table 3 shows the effects of patient characteristics on quality measures based on logistic regression for binary quality metrics. Quality metrics with extremely high or extremely low prevalence were not examined in the analysis. Age of patient was significantly associated ( $P < 0.05$ ) with QM1 (ERCP performed for an appropriate indication (OR 1.05 (1.03–1.07))), QM3 (pre-procedure history and directed physical examination performed and documented (OR 1.09 (1.04–1.14))), QM4 (risk for adverse events assessed and documented before sedation started (OR 1.03 (1.01–1.06))), and QM11 (precut sphincterotomy for cannulation (OR 1.21 (1.16–1.27))), meaning that adherence to these measures increased as patient age increased.

Sex of the patient was significantly associated with QM3 (pre-procedure history and directed physical examination performed and documented (OR 0.79 (0.67–0.93))), QM6 (deep cannulation of the ducts of interest is documented (OR 1.73 (1.05–2.83))), QM8 (pancreatic cannulation when not an intended target (OR 1.53 (1.30–1.80))), and QM9 (pancreatic injection when not an intended target (OR 1.51 (1.26–1.82))), with higher quality measures of QMs 3, X, Y for men and QMs, 6, 8, 9 for women.

Race (white versus other) was significantly associated with QM1 (ERCP performed for an appropriate indication (OR 1.11 (1.03–1.19))), QM3 (pre-procedure history and directed physical examination are performed and documented (OR 0.83 (0.69–0.99))), QM8 (pancreatic cannulation when not an intended target (OR 0.78 (0.66–0.92))) and QM9 (pancreatic injection when not an intended target (OR 0.77 (0.63–0.93))).

Figure 2 shows the unadjusted provider specific quality measurements with 95% confidence intervals (CI). Appendix Table 2 shows the adjusted and unadjusted proportions of quality measures by provider.

### Pre-Procedure Quality Measures

“Appropriate indication” ranged from 78.7–83.3% among the six providers with statistically significant variation ( $P < 0.0001$ ). Provider 6 fell below the 80% quality metric benchmark even with the 95% CI (78.7–79.9), results that persisted after adjustment for patient characteristics of age, gender, and race. “Informed consent documented” ranged from 99.4–100%. All providers met the recommended benchmark for this quality metric; no adjustment was done for patient factors given the high adherence rate and lack of variation. “Pre-

procedure H&P documented” ranged from 1.6–3.1%. All six providers failed to meet the performance benchmark with minimal change after adjustment for patient factors. There was statistical ( $P = P = 0.0195$ ) but not clinically significant difference among providers. “Adverse event risk documented” ranged among the six providers from 8.8–92.6% ( $P < 0.0001$ ). All six providers failed to meet the >98% benchmark with or without covariate adjustment. Provider 1 (92.6%) and Provider 6 (83.9%) were markedly different than Providers 2–5 (8.8–10.9%).

All providers averaged more than 100 ERCPs/year (QM5) with range from 282.7–570.9). Appendix Table 1 shows the breakdown of ERCP procedures within the dataset.

### Intra-Procedure Quality Measures

The documentation of deep cannulation showed limited clinical variation among the providers (97.2–99.4%). All six providers exceeded the >98% quality benchmark within their 95% CI. There was a statistically significant difference ( $P = 0.011$ ) among the six providers for this quality measure. “Unsuccessful stone removal < 1 cm” showed no variation (99.0–100%) with all six providers meeting the > 98% benchmark. “Pancreatic cannulation when not the intended target” ranged from 15.6–26.0% with clinically and statistically significant ( $P < 0.0001$ ) among the six providers. Adjustment for patient factors did not markedly change the proportions. The “pancreatic injection when not the intended target” ranged from 8.7–22.6%. This was both clinically and statistically significant ( $P < 0.0001$ ) among the six providers. Providers 1, 2, and 5 were clinically lower (8.7–11.8%) than Providers 3, 4, and 6 (18.9–22.6%). Adjustment for age, gender, and race did not markedly affect the metric.

Quality measure 10 (pancreatic stent placement if pancreatic duct cannulated and QM10<sup>^</sup> (addition of rectal indomethacin to QM10) ranged from 17.9–54.1% and 20.0–54.1% respectively. Both metrics showed statistically significant ( $P = 0.002$  and  $P = 0.0009$ ) variation among the six providers.

“Precut sphincterotomy for cannulation” ranged from 1.1–4%. There was significant variation among the six providers ( $P < 0.0001$ ). Adjustment for age, gender, and race did not markedly affect the metric.

### Post-Procedure Quality Measures

Perforation rate among the six providers ranged from 0.1–0.2% with a targeted quality measure of 0.2%. Significant bleeding occurred in 0.1–0.2% of patients within 7 days of the procedure, well below the 1% cut-off for procedural quality.

## DISCUSSION

ERCP is a challenging procedure with high risk for complications and technical failure. The risks increase exponentially when the procedure is performed for non-obstructive indications (e.g., sphincter of Oddi dysfunction or idiopathic acute pancreatitis). Feedback to colonoscopists on their adenoma detection rate (ADR) improves provider performance<sup>17</sup> and is associated with the subsequent risk of colorectal cancer<sup>5</sup>. It is plausible that feedback to

ERCP providers on their adherence to national recommendations would also improve the quality of ERCP services provided. The primary objective of this study was to develop a feasible method to track quality metrics in ERCP. An example of a stoplight report card is shown in Figure 3. Utilizing an existing, open-source based NLP system; we extracted quality measures over an 8.5-year period and compared them across individual providers and to society guidelines. This work is the first attempt to assess ERCP quality measures using NLP, and supports the feasibility of applying these techniques to larger datasets across multiple health care systems.

In our study we demonstrate clinically significant variation (e.g. rate of pancreatic cannulation when not the intended target) among providers, even in a highly skilled group of endoscopists at a single referral center (>9% variation). Given that one of the high risks for ERCP is development of post-ERCP pancreatitis, this may be a high-impact quality metric even among high-volume providers of ERCP. This knowledge may guide quality improvement projects to enhance appropriate documentation and identify providers who are not meeting society-endorsed benchmarks. While many of the measures (e.g. document pre-procedure H&P) may not reflect quality as these are often done externally to the report, they are contained within the quality tracking measurements for all endoscopic procedures.

The study has several limitations. First, the sample is restricted to ERCPs performed at a regional referral center. While this high volume unit does not reflect the general patient population, adherence to quality measures should apply to all ERCP providers. Furthermore, this study seeks to develop and validate a feasible method for assessing ERCP quality measures, not to report compliance with society guidelines.

A second limitation is that a single endoscopy software (Provation® MD; Wolters Kluwer) was utilized during the study period. This greatly enhances the ability for text mining and natural language processing to accurately detect specific concepts (e.g. 100% accuracy for providers). However, our group has shown that this technique can be applied to other institutions and accurately measure variables despite different methods for text document entry (e.g. dictation and endoscopy software)<sup>16</sup>. We also made assumptions about ICD-9-CM coding in relation to the procedure. This can be seen with the post-sphincterotomy bleeding rate allowing for 7 days after the procedure for any event associated with a specific ICD-9-CM code (285.1). With this assumption we may pick up non-ERCP related bleeding and/or bleeding not due to sphincterotomy. Additional ICD-9-CM coding such as 578.9 (GI bleeding) and 998.11 (hemorrhage complicating a procedure) might be utilized in the future to expand the identification of delayed complications.

Adjustment was done on multiple measures for this study; however, indication was not utilized for statistical adjustment and may significantly impact intra-procedure and post-procedure quality measures. Grouping ICD-9 CM codes into biliary, pancreatic, or dual indications may allow for adjustment in the future.

Last, when we compared our identification of ERCP documents by NLP to those produced by the endowriter software we noticed a difference of 4.1%. While this is a relatively small discrepancy, there are multiple potential explanations for this discrepancy. First, NLP may



not have correctly identified the documents as ERCP procedure reports from the 63,119 documents that were indexed as ERCP in the health information exchange. Second, the documents derived from the endowriter may not have been completely transmitted to the health information exchange, or have been duplicates. Last, the procedure report may have been filed under a different document type (e.g. operative report). Still, a discrepancy rate of 4.1% is low and would be unlikely to impact these observations.

## Conclusion

Overall, our study shows that NLP and data mining are capable of tracking adherence to ERCP-specific quality measures. Among six providers of ERCP at a single academic referral center there was significant variation in selected quality measures. The next step is to prove the external validity of this technique by assessing these measures using data from multiple institutions that utilize different types of electronic medical records.

## Acknowledgments

Dr. Timothy Imler had full access to all of the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis. The authors disclose that Drs. Imler has filed for provisional patent (IURTC-14098-01-US-E) for this work under the name Tracking Real-time Assessment of Quality Monitoring in Endoscopy (TRAQ-ME) through the Indiana University Research and Technology Corporation (IURTC). This work was performed at the Regenstrief Institute, Indianapolis, Indiana, and was supported in part by the American Society for Gastrointestinal Endoscopy Covidien Senior Investigator Mentoring Award (Imperiale) and the American Society for Gastrointestinal Endoscopy Career Development Award (Imler).

### Grant Support

This work was performed at the Regenstrief Institute, Indianapolis, Indiana, and was supported in part by the American Society for Gastrointestinal Endoscopy Covidien Senior Investigator Mentoring Award (Imperiale) and the American Society for Gastrointestinal Endoscopy Career Development Award (Imler).

## Abbreviations

<b>ERCP</b>	Endoscopic retrograde cholangiopancreatography
<b>NLP</b>	Natural language processing

## References

1. Adler, DG., Lieb, JG., II, Cohen, J., et al. Gastrointest Endosc. 2014. Quality indicators for ERCP.
2. Rex, DK., Schoenfeld, PS., Cohen, J., et al. Gastrointest Endosc. 2014. Quality indicators for colonoscopy.
3. Park, WG., Shaheen, NJ., Cohen, J., et al. Gastrointest Endosc. 2014. Quality indicators for EGD.
4. Wani, S., Wallace, MB., Cohen, J., et al. Gastrointest Endosc. 2014. Quality indicators for EUS.
5. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med.* 2014; 370:1298–306. [PubMed: 24693890]
6. Solad Y, Wang C, Laine L, et al. Influence of colonoscopy quality measures on patients' colonoscopist selection. *Am J Gastroenterol.* 2015; 110:215–9. [PubMed: 25070055]
7. Hewett DG, Rex DK. Improving colonoscopy quality through health-care payment reform. *Am J Gastroenterol.* 2010; 105:1925–33. [PubMed: 20551937]
8. Anderson MA, Fisher L, et al. Committee ASoP. Complications of ERCP. *Gastrointest Endosc.* 2012; 75:467–73. [PubMed: 22341094]

9. 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]
10. Cote GA, Imler TD, Xu H, et al. Lower provider volume is associated with higher failure rates for endoscopic retrograde cholangiopancreatography. *Med Care.* 2013; 51:1040–7. [PubMed: 24226304]
11. 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]
12. Imler TD, Morea J, Kahi C, et al. Natural language processing accurately categorizes findings from colonoscopy and pathology reports. *Clin Gastroenterol Hepatol.* 2013; 11:689–94. [PubMed: 23313839]
13. Mehrotra A, Dellon ES, Schoen RE, et al. Applying a natural language processing tool to electronic health records to assess performance on colonoscopy quality measures. *Gastrointest Endosc.* 2012; 75:1233–9. e14. [PubMed: 22482913]
14. Harkema H, Chapman WW, Saul M, et al. Developing a natural language processing application for measuring the quality of colonoscopy procedures. *J Am Med Inform Assoc.* 2011; 18(Suppl 1):i150–6. [PubMed: 21946240]
15. Gawron AJ, Thompson WK, Keswani RN, et al. Anatomic and Advanced Adenoma Detection Rates as Quality Metrics Determined via Natural Language Processing. *Am J Gastroenterol.* 2014; 109:1844–9. [PubMed: 24935271]
16. Imler TD, Morea J, Kahi C, et al. Multi-Center Colonoscopy Quality Measurement Utilizing Natural Language Processing. *Am J Gastroenterol.* 2015
17. Kahi CJ, Ballard D, Shah AS, et al. Impact of a quarterly report card on colonoscopy quality measures. *Gastrointest Endosc.* 2013; 77:925–31. [PubMed: 23472996]
18. Regenstrief Institute L. Health Information Exchange. 2014; 2014
19. Biondich PG, Grannis SJ. The Indiana network for patient care: an integrated clinical information system informed by over thirty years of experience. *J Public Health Manag Pract.* 2004; (Suppl):S81–6. [PubMed: 15643364]
20. 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]
21. Apache.org. UIMA. 2014.
22. Apache.org. License. 2014.
23. Apache.org. Solr. 2014.
24. ClinicalTrials.gov. Stent vs. Indomethacin for Preventing Post-ERCP Pancreatitis (SVI). 2016.
25. Imler TD, Morea J, Kahi C, et al. Multi-center colonoscopy quality measurement utilizing natural language processing. *Am J Gastroenterol.* 2015; 110:543–52. [PubMed: 25756240]

## Appendix

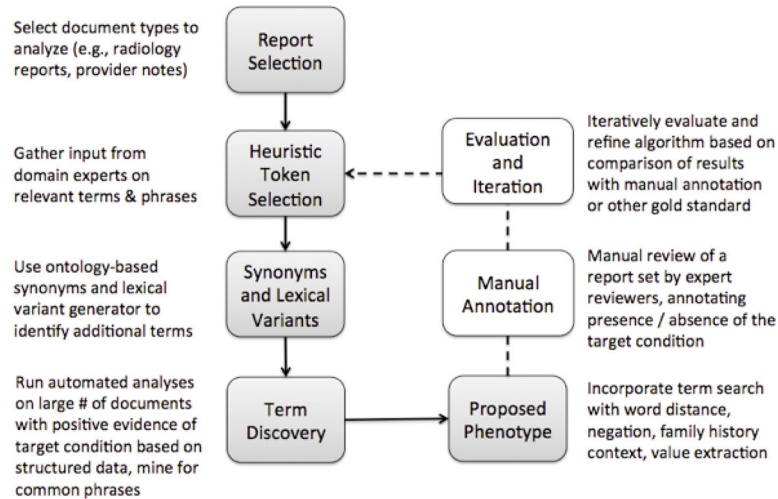
### ICD-9 CM Inclusion for Metric 1

Appropriate Indication codes included; calculus of bile duct (574.3\*, 574.4\*, 574.5\*), cholangitis (576.1), obstruction of bile duct (576.2), fistula of bile duct (576.4), spasm of sphincter of Oddi (576.5), other specified disorders of biliary tract (576.8), acute pancreatitis (577.1), 577.2 (cyst and pseudocyst of pancreas), 577.1 (chronic pancreatitis), 577.8 (other specified disease of pancreas), 751.7 (anomalies of pancreas), 157.\* (malignant neoplasm of pancreas), 156.\* (malignant neoplasm of gallbladder and extrahepatic bile ducts), and 155.\* (malignant neoplasm of liver and intrahepatic bile ducts). All codes were linked to an

appropriate ERCP related CPT code within the data set. The metric rate was number of appropriately identified ERCP/number of total ERCP by provider and in aggregate.

### ICD-9 CM Inclusion for Metric 6–9

Inclusion codes were calculus of bile duct (574.3\*, 574.4\*, 574.5\*) with exclusion codes of spasm of sphincter of Oddi (576.5), acute pancreatitis (577.1), 577.2 (cyst and pseudocyst of pancreas), 577.1 (chronic pancreatitis), 577.8 (other specified disease of pancreas), 751.7 (anomalies of pancreas), and 157.\* (malignant neoplasm of pancreas).



**Appendix Figure 1.**  
Overview of the phenotype development process for unstructured data

**Appendix Table 1**

Breakdown of ERCP procedures within dataset.

Provider	# of ERCP	Per Year Rate (Std)	# of Patients	Age (Std)	Female	White
Provider 1	1696	282.7 (113.3)	1060	53.7 (17.6)	991 (58.5%)	1266 (74.6%)
Provider 2	5133	570.4 (194.5)	3084	51.9 (18.0)	3174 (61.8%)	3946 (76.9%)
Provider 3	4455	495.3 (192.9)	3103	51.6 (17.0)	2812 (63.1%)	3460 (77.7%)
Provider 4	2680	297.9 (83.7)	1804	56.3 (17.3)	1497 (55.9%)	1998 (74.6%)
Provider 5	5138	570.9 (157.9)	2507	52.9 (18.1)	3025 (58.9%)	3920 (76.3%)
Provider 6	4572	508.0 (137.5)	2935	52.8 (17.9)	2614 (57.3%)	3386 (74.1%)
P Value	< 0.001		Not felt to be clinically significant			

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Appendix Table 2**

Adjusted and unadjusted proportions of quality measures by provider (excluding measure #5).

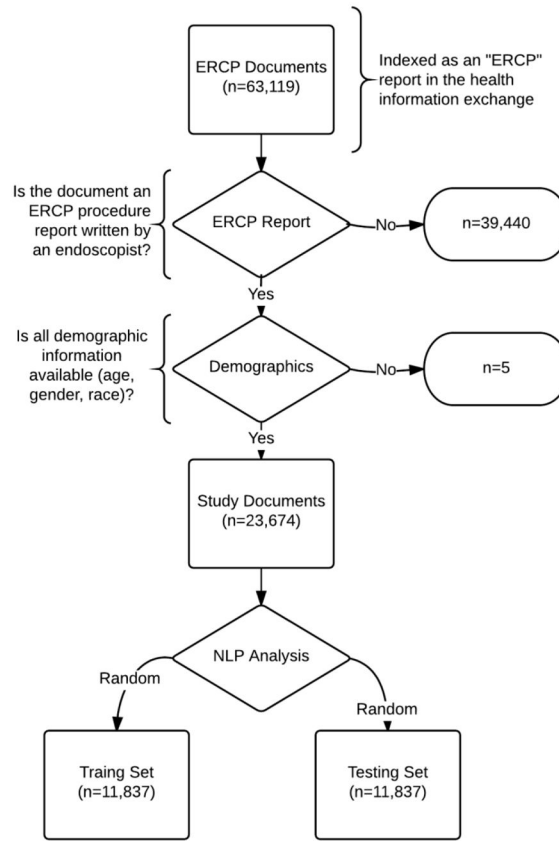
Quality Measure	Method	Provider	Numerator	Denominator	Mean % (95% CI)	P Value
Frequency with which endoscopy is performed for an appropriate indication (QM1)	Adjusted	Provider 1	.	.	79.5 (77.4–81.3)	<.0001
		Provider 2	.	.	79.6 (78.5–80.8)	
		Provider 3	.	.	80.7 (79.4–81.9)	
		Provider 4	.	.	78.8 (77.2–80.4)	
		Provider 5	.	.	83.0 (81.9–84.1)	
		Provider 6	.	.	78.4 (77.1–79.6)	
	Unadjusted	Provider 1	1354	1696	79.8 (77.9–81.7)	<.0001
		Provider 2	4099	5134	79.8 (78.7–80.9)	
		Provider 3	3603	4458	80.8 (79.7–82.0)	
		Provider 4	2129	2681	79.4 (77.9–80.9)	
		Provider 5	4280	5138	83.3 (82.3–84.3)	
		Provider 6	3600	4572	78.7 (77.6–79.9)	
Frequency with which informed consent is obtained and fully documented (QM2)	Unadjusted	Provider 1	1696	1696	100 (99.8–100)	*
		Provider 2	5134	5134	100 (99.9–100)	
		Provider 3	4432	4458	99.4 (99.1–99.6)	
		Provider 4	2679	2681	99.9 (99.7–100)	
		Provider 5	5137	5138	100 (99.9–100)	
		Provider 6	4570	4572	100 (99.8–100)	
Frequency with which pre- procedure history and directed physical examination are performed and documented (QM3)	Adjusted	Provider 1	.	.	1.7 (1.2–2.4)	0.0223
		Provider 2	.	.	2.9 (2.4–3.4)	
		Provider 3	.	.	3.3 (2.8–3.9)	
		Provider 4	.	.	3.0 (2.4–3.7)	
		Provider 5	.	.	2.5 (2.1–3.0)	
		Provider 6	.	.	2.8 (2.3–3.3)	
	Unadjusted	Provider 1	27	1696	1.6 (1.0–2.2)	0.0195
		Provider 2	139	5134	2.7 (2.3–3.2)	
		Provider 3	139	4458	3.1 (2.6–3.6)	
		Provider 4	79	2681	2.9 (2.3–3.6)	
		Provider 5	123	5138	2.4 (2.0–2.8)	
		Provider 6	120	4572	2.6 (2.2–3.1)	
Frequency with which risk for adverse events is assessed and documented before sedation is started (QM4)	Adjusted	Provider 1	.	.	92.5 (91.2–93.7)	<.0001
		Provider 2	.	.	9.0 (8.2–9.8)	
		Provider 3	.	.	10.3 (9.4–11.3)	
		Provider 4	.	.	10.8 (9.6–12.0)	
		Provider 5	.	.	8.8 (8.0–9.7)	

Quality Measure	Method	Provider	Numerator	Denominator	Mean % (95% CI)	P Value
	Unadjusted	Provider 6	.	.	83.9 (82.8–85.0)	<.0001
		Provider 1	1570	1696	92.6 (91.3–93.8)	
		Provider 2	459	5134	8.9 (8.2–9.7)	
		Provider 3	460	4458	10.3 (9.4–11.2)	
		Provider 4	292	2681	10.9 (9.7–12.1)	
		Provider 5	452	5138	8.8 (8.0–9.6)	
		Provider 6	3835	4572	83.9 (82.8–84.9)	
Frequency with which deep cannulation of the ducts of interest is documented (QM6)	Adjusted	Provider 1	.	.	97.2 (94.1–98.7)	0.0219
		Provider 2	.	.	99.4 (98.6–99.7)	
		Provider 3	.	.	97.1 (95.3–98.3)	
		Provider 4	.	.	99.1 (97.4–99.7)	
		Provider 5	.	.	98.2 (97.0–98.9)	
		Provider 6	.	.	98.0 (96.9–98.8)	
	Unadjusted	Provider 1	225	231	97.4 (94.4–99.0)	0.0107
		Provider 2	892	897	99.4 (98.7–99.8)	
		Provider 3	516	531	97.2 (95.4–98.4)	
		Provider 4	378	381	99.2 (97.7–99.8)	
		Provider 5	835	850	98.2 (97.1–99.0)	
		Provider 6	905	923	98.0 (96.9–98.8)	
Frequency with which common bile duct stones <1 cm in patients with normal bile duct anatomy are extracted successfully and documented (QM7)	Unadjusted	Provider 1	229	231	99.1 (96.9–99.9)	*
		Provider 2	896	897	99.9 (99.4–100)	
		Provider 3	526	531	99.1 (97.8–99.7)	
		Provider 4	377	381	99.0 (97.3–99.7)	
		Provider 5	847	850	99.6 (99.0–99.9)	
		Provider 6	923	923	100 (99.6–100)	
Frequency of pancreatic cannulation when not an intended target (QM8)	Adjusted	Provider 1	.	.	15.7 (11.5–21%)	<.0001
		Provider 2	.	.	16.9 (14.5–19.5)	
		Provider 3	.	.	27.7 (24.0–31.8)	
		Provider 4	.	.	26.0 (21.8–30.8)	
		Provider 5	.	.	20.3 (17.6–23.2)	
		Provider 6	.	.	24.5 (21.7–27.5)	
	Unadjusted	Provider 1	36	231	15.6 (10.9–20.3)	<.0001
		Provider 2	148	897	16.5 (14.1–18.9)	
		Provider 3	145	531	27.3 (23.5–31.1)	
		Provider 4	99	381	26.0 (21.6–30.4)	
		Provider 5	170	850	20.0 (17.3–22.7)	
		Provider 6	220	923	23.8 (21.1–26.6)	
Frequency of pancreatic injection when not an intended target (QM9)	Adjusted	Provider 1	.	.	8.8 (5.7–13.2)	<.0001
		Provider 2	.	.	12.1 (10.1–14.5)	

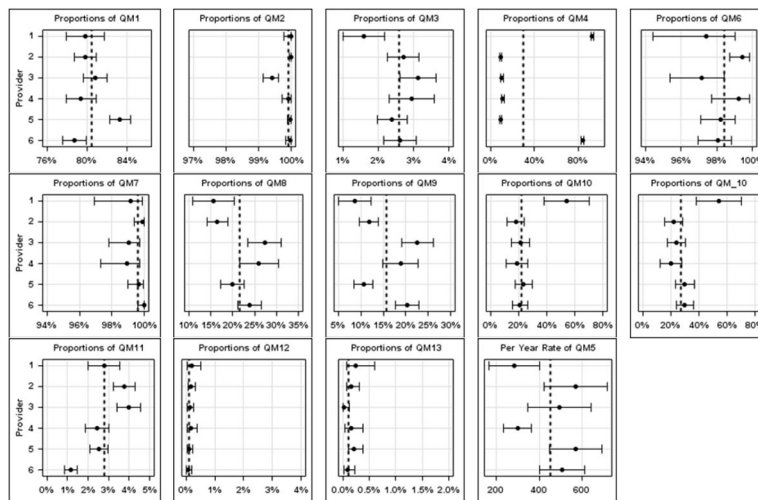
Quality Measure	Method	Provider	Numerator	Denominator	Mean % (95% CI)	P Value	
		Provider 3	.	.	22.9 (19.4–26.8)	<.0001	
		Provider 4	.	.	19.0 (15.3–23.4)		
		Provider 5	.	.	10.8 (8.8–13.1)		
		Provider 6	.	.	20.8 (18.2–23.7)		
	Unadjusted	Provider 1	20	231	8.7 (5.0–12.3)		
		Provider 2	106	897	11.8 (9.7–13.9)		
		Provider 3	120	531	22.6 (19.0–26.2)		
		Provider 4	72	381	18.9 (15.0–22.8)		
		Provider 5	90	850	10.6 (8.5–12.7)		
		Provider 6	187	923	20.3 (17.7–22.9)		
Frequency of pancreatic stent placement if pancreatic duct cannulated (QM10)	Adjusted	Provider 1	.	.	52.7 (36.7–68.2)	0.0012	
		Provider 2	.	.	17.8 (12.4–24.8)		
		Provider 3	.	.	21.0 (15.2–28.3)		
		Provider 4	.	.	18.7 (12.2–27.6)		
		Provider 5	.	.	23.4 (17.5–30.5)		
		Provider 6	.	.	20.7 (15.7–26.6)		
	Unadjusted	Provider 1	20	37	54.1 (38.0–70.1)		0.0002
		Provider 2	27	151	17.9 (11.8–24.0)		
		Provider 3	34	160	21.3 (14.9–27.6)		
		Provider 4	19	100	19.0 (11.3–26.7)		
		Provider 5	40	170	23.5 (17.2–29.9)		
		Provider 6	46	221	20.8 (15.5–26.2)		
Frequency of pancreatic stent placement and/or indomethacin given if pancreatic duct cannulated (QM10 <sup>^</sup> )	Adjusted	Provider 1	.	.	53.1 (37.1–68.5)	0.0020	
		Provider 2	.	.	21.5 (15.6–28.8)		
		Provider 3	.	.	23.1 (17.1–30.5)		
		Provider 4	.	.	19.8 (13.1–28.8)		
		Provider 5	.	.	29.5 (23.1–37.0)		
		Provider 6	.	.	29.4 (23.6–35.9)		
	Unadjusted	Provider 1	20	37	54.1 (38.0–70.1)		0.0009
		Provider 2	33	151	21.9 (15.3–28.4)		
		Provider 3	38	160	23.8 (17.2–30.3)		
		Provider 4	20	100	20.0 (12.2–27.8)		
		Provider 5	51	170	30.0 (23.1–36.9)		
		Provider 6	66	221	29.9 (23.8–35.9)		
Frequency of precut sphincterotomy for cannulation	Adjusted	Provider 1	.	.	2.6 (2.0–3.5)	<.0001	
		Provider 2	.	.	3.6 (3.1–4.2)		
		Provider 3	.	.	3.9 (3.3–4.5)		
		Provider 4	.	.	2.2 (1.7–2.8)		
		Provider 5	.	.	2.4 (2.0–2.9)		

Quality Measure	Method	Provider	Numerator	Denominator	Mean % (95% CI)	P Value
	Unadjusted	Provider 6	.	.	1.1 (0.8–1.4)	<.0001
		Provider 1	47	1696	2.8 (2.0–3.6)	
		Provider 2	193	5134	3.8 (3.2–4.3)	
		Provider 3	177	4458	4.0 (3.4–4.5)	
		Provider 4	65	2681	2.4 (1.8–3.0)	
		Provider 5	130	5138	2.5 (2.1–3.0)	
		Provider 6	52	4572	1.1 (0.8–1.4)	
Frequency of perforation due to ERCP (within 7 days)	Unadjusted	Provider 1	3	1696	0.2 (0.0–0.5)	*
		Provider 2	8	5134	0.2 (0.1–0.3)	
		Provider 3	5	4458	0.1 (0.0–0.3)	
		Provider 4	4	2681	0.1 (0.0–0.4)	
		Provider 5	5	5138	0.1 (0.0–0.2)	
		Provider 6	3	4572	0.1 (0.0–0.2)	
Rate of clinically significant hemorrhage after ERCP with or without sphincterotomy (within 7 days)	Unadjusted	Provider 1	4	1696	0.2 (0.1–0.6)	*
		Provider 2	8	5134	0.2 (0.1–0.3)	
		Provider 3	1	4458	0.0 (0.0–0.1)	
		Provider 4	4	2681	0.1 (0.0–0.4)	
		Provider 5	11	5138	0.2 (0.1–0.4)	
		Provider 6	4	4572	0.1 (0.0–0.2)	



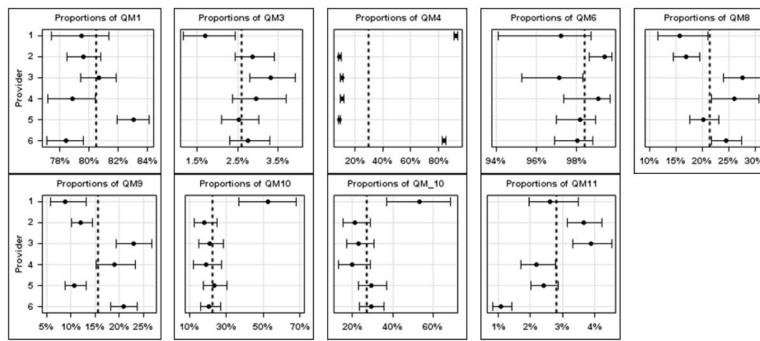


**Figure 1.**  
Flow chart for study.



**Figure 2.** Unadjusted provider specific quality measurements for ERCP with 95% confidence intervals.

- QM1 = endoscopy is performed for an appropriate indication
- QM2 = informed consent is obtained and fully documented
- QM3 = pre-procedure history and directed physical examination are performed and documented
- QM4 = risk for adverse events is assessed and documented before sedation is started
- QM5 = Volume of ERCs performed per year by endoscopist
- QM6 = deep cannulation of the ducts of interest is documented\*
- QM7 = common bile duct stones <1 cm in patients with normal bile duct anatomy are extracted successfully and documented
- QM8 = pancreatic cannulation when not an intended target\*
- QM9 = pancreatic injection when not an intended target\*
- QM10 = pancreatic stent placement if pancreatic duct cannulated
- QM11 = precut sphincterotomy for cannulation
- QM12 = perforation due to ERCP (within 7 days)
- QM13 = Rate of clinically significant hemorrhage after ERCP with or without sphincterotomy (within 7 days)



**Figure 3.** Adjusted provider specific quality measurements for ERCP with 95% confidence intervals.

- QM1 = endoscopy is performed for an appropriate indication
- QM3 = pre-procedure history and directed physical examination are performed and documented
- QM4 = risk for adverse events is assessed and documented before sedation is started
- QM6 = deep cannulation of the ducts of interest is documented
- QM8 = pancreatic cannulation when not an intended target
- QM9 = pancreatic injection when not an intended target
- QM10 = pancreatic stent placement if pancreatic duct cannulated
- QM11 = precut sphincterotomy for cannulation

**Provider, One Stoplight Report**

	Yearly		Quarterly			
	Current YTD	Previous Year	Qtr FY201x	Qtr FY201x	Qtr FY201x	Qtr FY201x
<b>Pre-Procedure</b>						
Frequency with which endoscopy is performed for an appropriate indication	79.5	x	x	x	x	x
Frequency with which informed consent is obtained and fully documented	100	x	x	x	x	x
Frequency with which pre-procedure history and directed physical examination are performed and documented	1.7	x	x	x	x	x
Frequency with which risk for adverse events is assessed and documented before sedation is started	92.5	x	x	x	x	x
Volume of ERCPs performed	1696	x	x	x	x	x
<b>Intra-Procedure</b>						
Frequency with which deep cannulation of the ducts of interest is documented	97.2	x	x	x	x	x
Frequency with which common bile duct stones <1 cm in patients with normal bile duct anatomy are extracted successfully and documented	99.1	x	x	x	x	x
Frequency of pancreatic cannulation when not an intended target	15.7	x	x	x	x	x
Frequency of pancreatic	8.8	x	x	x	x	x

stent placement if pancreatic duct cannulated						
Frequency of pancreatic stent placement if pancreatic duct cannulated	52.7	x	x	x	x	x
Frequency of precut sphincterotomy for cannulation	2.6	x	x	x	x	x
<b>Post-Procedure</b>						
Frequency of perforation due to ERCP (within 7 days)	0.2	x	x	x	x	x
Rate of clinically significant hemorrhage after ERCP with or without sphincterotomy (within 7 days)	0.2	x	x	x	x	x

**Figure 4.** Example provider quality report card based on Provider 1 overall measure. Green is for being above the benchmark including 95% CI. Yellow includes the benchmark in the 95% CI. Red is below the benchmark including the 95% CI.

**Table 1**

Quality measures for ERCP and method of extraction.

Metric #	Quality Indicator	Grade of Recommendation per Society Guidelines <sup>1</sup>	Performance Target	Extraction Method
<b>Pre-procedure</b>				
1	Endoscopy is performed for an appropriate indication	IC+	> 80%	ICD-9 CM
2	Informed consent is obtained and fully documented	3	> 98%	NLP
3	Pre-procedure history and directed physical examination are performed and documented	3	> 98%	NLP
4	Risk for adverse events is assessed and documented before sedation is started	3	> 98%	NLP
5	Volume of ERCPPs performed per year by endoscopist	IC	> 100	NLP
<b>Intra-procedure</b>				
6	deep cannulation of the ducts of interest is documented <sup>*</sup>	IC	> 98%	NLP and ICD-9 CM
7	Common bile duct stones <1 cm in patients with normal bile duct anatomy are extracted successfully and documented <sup>*</sup>	IC	90%	NLP and ICD-9 CM
8§	Pancreatic cannulation when not an intended target <sup>*</sup>	n/a	n/a	NLP and ICD-9 CM
9§	Pancreatic injection when not an intended target <sup>*</sup>	n/a	n/a	NLP and ICD-9 CM
10§	Pancreatic stent placement <i>if pancreatic duct cannulated</i>	n/a	n/a	NLP
11§	Pre-cut sphincterotomy for cannulation	n/a	n/a	NLP
<b>Post-procedure</b>				
12	perforation due to ERCP (within 7 days)	2C	0.2	ICD-9 CM
13	Rate of clinically significant hemorrhage after ERCP with or without sphincterotomy (within 7 days)	IC	< 1	ICD-9 CM

<sup>\*</sup> For indication of choledocholithiasis (ICD-9 CM 574.3\*, 574.4\*, 574.5\*) with exclusions of sphincter of Oddi dysfunction or pancreatic pathology by ICD-9 CM and NLP.

<sup>§</sup> Measures added by investigators as potential quality metrics.

**Table 2**

Validation metrics for natural language processing on ERCP quality measures.

Measure	True Positive (n=50)	True Negative (n=50)	Testing Set Accuracy %	Testing Set Precision
ERCP Procedure Report	50	50	100	1
Provider 1 as Provider	50	50	100	1
Provider 2 as Provider	50	50	100	1
Provider 3 as Provider	50	50	100	1
Provider 4 as Provider	50	50	100	1
Provider 5 as Provider	50	50	100	1
Provider 6 as Provider	50	50	100	1
<b>Pre-procedure</b>				
endoscopy is performed for an appropriate indication (QM1)	n/a	n/a	n/a	n/a
informed consent is obtained and fully documented (QM2)	50	11*	100	1
pre-procedure history and directed physical examination are performed and documented (QM3)	50	50	100	1
risk for adverse events is assessed and documented before sedation is started (QM4)	50	50	100	1
Volume of ERCPs performed per year by endoscopist (QM5)	n/a	n/a	n/a	n/a
<b>Intra-procedure</b>				
deep cannulation of the ducts of interest is documented (QM6)	50	47	97	1
common bile duct stones <1 cm in patients with normal bile duct anatomy are extracted successfully and documented (QM7)	49	4**	96.4	98
pancreatic cannulation when not an intended target (QM8)	48	50	98	96
pancreatic injection when not an intended target (QM9)	48	46	94	96
pancreatic stent placement if pancreatic duct cannulated (QM10)	46	44	90	92
precut sphincterotomy for cannulation (QM11)	42	49	91	84
<b>Post-procedure</b>				
perforation due to ERCP (within 7 days) (QM12)	n/a	n/a	n/a	n/a
Rate of clinically significant hemorrhage after ERCP with or without sphincterotomy (within 7 days) (QM13)	n/a	n/a	n/a	n/a

\* Only 11 reports available for review for the negated search.

\*\* Only 5 reports available for review for the negated search.

**Table 3**

Effects of patient characteristics on quality measurements.

Quality Measure	Covariates	Odds Ratio	Standard Error	Lower Confidence Limit	Upper Confidence Limit	P Value
Endoscopy is performed for an appropriate indication (QM1)	10-Year Increase in Age	1.05	0.01	1.03	1.07	< 0.01
	Female	0.95	0.03	0.89	1.01	0.12
	White vs Other	1.11	0.04	1.03	1.19	0.01
Pre-procedure history and directed physical examination are performed and documented (QM3)	10-Year Increase in Age	1.09	0.03	1.04	1.14	< 0.01
	Female	0.79	0.06	0.67	0.93	< 0.01
	White vs Other	0.83	0.08	0.69	0.99	0.04
Risk for adverse events is assessed and documented before sedation is started (QM4)	10-Year Increase in Age	1.03	0.01	1.01	1.06	0.01
	Female	0.98	0.04	0.90	1.06	0.59
	White vs Other	1.00	0.05	0.91	1.10	0.99
Deep cannulation of the ducts of interest is documented (QM6)	10-Year Increase in Age	1.07	0.07	0.93	1.22	0.34
	Female	1.73	0.44	1.05	2.83	0.03
	White vs Other	0.90	0.26	0.51	1.58	0.72
Pancreatic cannulation when not an intended target (QM8)	10-Year Increase in Age	1.00	0.02	0.96	1.05	0.82
	Female	1.53	0.13	1.30	1.80	0.00
	White vs Other	0.78	0.07	0.66	0.92	0.00
Pancreatic injection when not an intended target (QM9)	10-Year Increase in Age	0.98	0.02	0.94	1.03	0.48
	Female	1.51	0.14	1.26	1.82	0.00
	White vs Other	0.77	0.08	0.63	0.93	0.01
Pancreatic stent placement if pancreatic duct cannulated (QM10)	10-Year Increase in Age	1.05	0.04	0.96	1.13	0.29
	Female	1.15	0.21	0.81	1.63	0.44
	White vs Other	1.04	0.19	0.73	1.49	0.82
QM10 with addition of indomethacin use	10-Year Increase in Age	1.00	0.04	0.92	1.08	0.93
	Female	1.12	0.19	0.81	1.55	0.49
	White vs Other	1.12	0.19	0.80	1.57	0.51
Precut sphincterotomy for cannulation (QM11)	10-Year Increase in Age	1.21	0.03	1.16	1.27	0.00
	Female	0.98	0.08	0.84	1.15	0.79
	White vs Other	1.01	0.10	0.84	1.22	0.88