



Measuring Surgical Quality: A National Clinical Registry Versus Administrative Claims Data

Laura M. Enomoto MD¹, Christopher S. Hollenbeak PhD^{1,2}, Neil H. Bhayani MHS MD¹, Peter W. Dillon MS MD¹, Niraj J. Gusani MS MD^{1,2}

¹ Department of Surgery, Penn State Milton S. Hershey Medical Center, Hershey, PA

² Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA

Introduction

Administrative data are readily available and inexpensive but its accuracy and reliability in assessing surgical outcomes has repeatedly been questioned. This study compared post-operative complications of patients who underwent pancreaticoduodenectomy (PD) recorded in the National Surgical Quality Improvement Program (NSQIP), a risk-adjusted, validated clinical database to patients who underwent PD recorded in the Healthcare Cost and Utilization Project (HCUP) National Inpatient Sample (NIS), a widely utilized administrative dataset.

Methods

This was a retrospective cohort study that included patients who underwent PD recorded in NSQIP and NIS from 2005-2010. PD was identified in NSQIP using CPT codes (48150, 48152, 48153, and 48154) and in NIS using the ICD-9 primary procedure code for PD (52.7).

In NIS, 18 post-operative adverse events were identified using ICD-9 codes. These corresponded to 17 post-operative NSQIP variables. These variables were studied as post-operative outcomes using univariate analysis followed by logistic or generalized linear regression.

Results

There were 8,800 PDs recorded in NSQIP and 9,830 PDs recorded in NIS after excluding 57 patients due to missing covariates or outcome variables. Patients were well matched in age and sex. Distributions of patient race and admission source were similar between the two databases.

Using univariate analysis, patients undergoing PD recorded in NSQIP demonstrated a significantly higher incidence of systemic sepsis, wound

infection, bleeding requiring transfusion, deep vein thrombosis, pulmonary embolism, cardiac arrest, and cerebrovascular accident.

Table 1 provides the results of logistic and generalized linear regression models of each post-operative event controlling for covariates and using NSQIP as the reference database. Relative to patients in NSQIP, patients in NIS were more likely to have poorer outcomes, and nearly all of the effects were significant.

Discussion

Our study demonstrated that risk of most post-operative complications was significantly higher in patients undergoing PD recorded in NIS compared to patients undergoing PD recorded in NSQIP, which is

in contrast to multiple studies which demonstrated a higher rate of post-operative complications recorded in clinical databases compared to administrative claims data.¹⁻³ Unlike previous studies which included a diverse case mix, however, we specifically analyzed PD, a complex procedure known to be highly morbid with a longer length of stay than less complex general surgery cases. NSQIP includes patient outcomes for 30 days post-operatively, usually including some time after discharge from the hospital, while NIS only includes outcomes collected in the hospital setting during the index admission. Because of the longer length of stay following PD than most general surgery cases, patients who underwent PD recorded in NIS could have more complications captured during their

index admission than studies that included procedures with shorter lengths of stay where complications could occur after discharge and thus not be captured in NIS.

Despite well-matched patients, there is still considerable discordance between NSQIP and NIS in the assessment of post-operative complications following PD. Although using administrative data to measure clinical outcomes is appealing due to its wide availability and low cost, these findings suggest that use of ICD-9 codes to adjust for patient characteristics and post-operative complications delivers significantly different results than a validated clinical database. This underscores the value of recognizing the capabilities and limitations of each data source. Administrative data was not originally designed to measure patient outcomes, and should therefore not be expected to accurately and reliably deliver them.

References

1. Best WR, Khuri SF, Phelan M, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: results from the Department of Veterans Affairs National Surgical Quality Improvement Program. *JACS*. Mar 2002;194(3):257-266.
2. Cima RR, Lackore KA, Nehring SA, et al. How best to measure surgical quality? Comparison of the Agency for Healthcare Research and Quality Patient Safety Indicators (AHRQ-PSI) and the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) postoperative adverse events at a single institution. *Surgery*. Nov 2011;150(5):943-949.
3. Romano PS, Schembri ME, Rainwater JA. Can administrative data be used to ascertain clinically significant postoperative complications? *Amer J of Med Quality: Official Journal of the American College of Medical Quality*. Jul-Aug 2002;17(4):145-154.

Table 1. Results of logistic regression models of effect of type of database on post-operative complication, controlling for other covariates. There were no patients with bleeding requiring transfusion or pulmonary emboli recorded in NIS, thus odds ratios were not calculated for these two post-operative events. NSQIP is the reference database for all odds ratios.

Variable	Odds Ratio	95% Confidence Interval		P-Value
		Lower	Upper	
Systemic sepsis	0.38	0.32	0.45	< 0.001
Superficial or deep wound infection	0.91	0.79	1.06	0.235
Deep venous thrombosis	0.18	0.10	0.33	< 0.001
Cardiac arrest with CPR	0.45	0.26	0.75	0.002
Cerebrovascular accident	0.04	0.00	0.34	0.003
Fail to wean from ventilator	1.36	1.14	1.62	0.001
Unplanned intubation	1.30	1.08	1.56	0.005
Pneumonia	1.51	1.26	1.81	< 0.001
Urinary tract infection	1.42	1.18	1.71	< 0.001
Coma	12.09	6.90	21.19	< 0.001
Wound dehiscence	1.00	0.72	1.39	0.979
Renal failure	1.67	1.17	2.38	0.005
Renal insufficiency	2.39	1.63	3.50	< 0.001
Myocardial infarction	1.75	1.06	2.91	0.030
Graft/prosthetic failure	9.35	4.34	20.14	< 0.001
Length of Stay	1.10*	0.56	1.64	<0.001

* Marginal Effect