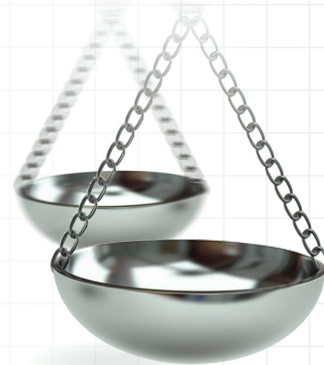




Using a society database to evaluate a patient safety collaborative: the Cardiovascular Surgical Translational Study

Journal of **Comparative Effectiveness Research**

Yea-Jen Hsu¹, Andrzej S Kosinski², Amelia S Wallace², Paramita Saha-Chaudhuri³, Bickey H Chang⁴, Kathleen Speck⁴, Michael A Rosen^{4,5}, Ayse P Gurses^{1,4,5}, Anping Xie^{4,5}, Shu Huang⁴, Duke E Cameron⁶, David A Thompson^{1,4,5,7} & Jill A Marsteller^{*,1,4,5}

¹Department of Health Policy & Management, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Baltimore, MD 21205, USA

²Duke Clinical Research Institute, Duke University School of Medicine, 2400 Pratt Street, Durham, NC 27705, USA

³Department of Epidemiology Biostatistics & Occupational Health, McGill University, 1020 Pine Avenue, West Montreal, Quebec, Canada

⁴Armstrong Institute for Patient Safety & Quality, Johns Hopkins Medicine, 750 E. Pratt Street, Baltimore, MD 21202, USA

⁵Department of Anesthesiology & Critical Care Medicine, Johns Hopkins School of Medicine, 1800 Orleans Street, Baltimore, MD 21287, USA

⁶Division of Cardiac Surgery, Johns Hopkins School of Medicine, 1800 Orleans Street, Baltimore, MD 21287, USA

⁷Division of Acute & Chronic Care, Johns Hopkins School of Nursing, 525 N. Wolfe Street, Baltimore, MD 21205, USA

*Author for correspondence: Tel.: +1 410 614 2602; Fax: +1 410 955 3249; jmarste2@jhu.edu

Aim: To assess the utility of using external databases for quality improvement (QI) evaluations in the context of an innovative QI collaborative aimed to reduce three infections and improve patient safety across the cardiac surgery service line. **Methods:** We compared changes in each outcome between 15 intervention hospitals (infection reduction protocols plus safety culture intervention) and 52 propensity score-matched hospitals (feedback only). **Results:** Improvement trends in several outcomes among the intervention hospitals were not statistically different from those in comparison hospitals. **Conclusion:** Using external databases such as those of professional societies may permit comparative effectiveness assessment by providing concurrent comparison groups, additional outcome measures and longer follow-up. This can better inform evaluation of continuous QI in healthcare organizations.

First draft submitted: 1 June 2018; Accepted for publication: 3 October 2018; Published online: 11 December 2018

Keywords: cardiac surgery • hospital-acquired infections • patient safety • program evaluation • quality improvement • secondary data analysis • study design

Evaluators of quality improvement and patient safety (QI/PS) programs examine the impact of interventions on process measures and patient outcomes, and sometimes the context of the program, seeking to close the gaps between evidence-based practice and real-world practice and to inform the development of best implementation and dissemination strategies. However, many QI/PS projects and their evaluation efforts, due to capacity constraints, are launched with a less rigorous evaluation design or are based on anecdotal and intuitive accounts, which makes it difficult to provide adequate and meaningful evidence for stakeholders to make decisions [1,2]. For example, the majority of the QI/PS evaluation studies published in peer-reviewed journals rely on simple pre-post designs and lack a concurrent comparison group [3]. The uncontrolled pre-post evaluation design is generally more prone to bias introduced by secular trends or external changeable factors, which poses challenges for attributing observed changes to the intervention in these studies [4]. Moreover, the evaluations often only collect outcome measures on which the intervention is specifically hypothesized to have an impact (which may miss similar trends in unrelated variables [suggesting omitted variable bias], or alternatively, unintended consequences) [5]. Changes in process measures are often not assessed, not reported, or have substantial missing data. Sustainability of these projects is generally not followed or reported [6].

Evaluation experts have urged QI/PS evaluators to use stronger evaluation designs, such as multiple observations before and after the intervention (i.e., time-series design) or to look for opportunities for a trial design with a comparison group when a randomized controlled trial is not feasible [1,4]. Pre-existing databases, such as those of professional societies, may have rich data that can be used for the purpose of QI/PS evaluations. With consistent and rigorous data collection protocols and a large group of members, these databases may allow the use of concurrent comparison groups, additional impact measures and longer follow-up periods to evaluate the effectiveness of the QI/PS interventions with appropriate control groups and to assist stakeholders in making decisions with informative evidence.

The goal of this study is to use the Society of Thoracic Surgeons Adult Cardiac Surgery Database (STS ACSDB) – a mature database to which most hospitals performing cardiac surgery in the US report data (www.sts.org/sections/stsnationaldatabase/) – to evaluate the Cardiovascular Surgical Translational Study (CSTS). This patient safety research program aimed to reduce hospital-acquired infections (HAIs) and improve patient safety across cardiac surgery units within a hospital, from operating rooms (ORs) and intensive care units (ICUs) to inpatient floor units [7]. The CSTS was designed to extend successful patient safety efforts, such as central line-associated bloodstream infections in ICUs [8–10], beyond individual clinical settings to an entire service line. To our knowledge, few initiatives have sought to improve patient safety and prevent HAIs for an entire service line where intrahospital transitions of patients across units occur [11]. Among other issues, transitions and handoffs of patients pose significant risks to the safety of hospitalized patients, particularly those with complex needs, and result in hospital complications and longer lengths of stay [12–14]. Literature suggests the importance of facilitating effective communication and teamwork across hospital units to help bridge the gaps in cross-unit transitions and handoffs, such as through refining and standardizing handoff protocols [15–20]. The multifaceted intervention in the CSTS included evidence-based infection prevention bundles and an adaptive component to create a positive safety culture and commitment for all the patient safety efforts. The project targeted three types of infections, including central line-associated bloodstream infection (CLABSI), surgical site infection (SSI) and ventilator-associated pneumonia (VAP), with a focus on interventions preventing new infection type added every 4–6 months. The study design was a nonrandomized pre-post evaluation without comparisons, using hospital-reported data collected throughout the project. This design was selected to maximize the dissemination of a proven intervention among participating sites, while examining the feasibility of targeting multiple infection types concurrently.

The hospital-reported data showed a reduction in both SSIs and CLABSIs after 2 years as compared with the baseline [7]. However, the reductions in each case were followed by a rebound in the third year, the sustainment period. The results of the VAP intervention were inconclusive due to the definition change from the Center for Disease Control and Prevention (CDC) in the middle of the study period and lack of baseline data that were consistent with the new definition. Moreover, the results from the Hospital Survey on Patient Safety suggested improved providers' perceptions regarding specific domains of patient safety culture, including teamwork across hospital units, handoffs, leadership and nonpunitive response to error; but decreased ratings regarding staffing, feedback and communication about errors, and communication openness. We hoped to learn more about the study's effects by examining information from a pre-existing, ongoing data collection external to the CSTS study.

The Society of Thoracic Surgeons' approach to improvement is to provide passive feedback of performance and benchmark data to its member hospitals, which receive STS reports but may or may not react to them. This study is therefore a comparative effectiveness examination of the CSTS interventions plus active performance feedback relative to audit and feedback only. We hypothesize that the CSTS hospitals will have greater improvement in assessed outcomes as compared with the matched comparison hospitals.

Materials & methods

The Cardiovascular Surgical Translational Study

The CSTS initially recruited 15 participating hospitals across 11 states in the USA. Four hospitals did not complete Institutional Review Board (IRB) approval from their own institution and dropped out at the earliest stage of the program. One hospital dropped out during the last year (sustainability year) due to lack of resources to support continuation of the project. Thus, ten hospitals participated in the CSTS from beginning to end.

The types of units that were included in the project consisted of ORs, ICUs, floor units and universal-bed units (a care model that manages patient care in one setting from postsurgery to discharge). To participate in the project, each hospital was required to assign an overall project manager (typically a nurse), and to assemble a QI team in each unit, usually composed of a physician and a nurse. The project used a multisite learning collaborative to train

the unit-specific QI teams by providing monthly conference calls and annual face-to-face conferences. Then, the QI teams educated other staff and implemented the interventions in their units.

All units were instructed to implement the Comprehensive Unit-based Safety Program (CUSP), a 5-step program designed to improve teamwork, communication and patient safety culture. The five steps were to educate staff on the science of improving patient safety with an educational video; identify hazards in the unit using the Staff Safety Assessment, asking how the next patient might be harmed; partner with a senior executive to help set priorities and provide resources; learn from one defect and implement improvements each month based on the set priorities; and implement teamwork tools that address individual unit's specific challenges regarding interprofessional relationships and work processes [21–23].

In addition to CUSP, the central research team developed evidence-based prevention bundles to help participating units set up protocols for infection reduction. The infection prevention bundles were introduced with a phased approach. The first training series on the CLABSI bundle was started in June 2011, followed by the SSI bundle which began in December 2011 and the VAP bundle in April 2012. The initial phase of implementation and evaluation was funded by the Agency for Healthcare Research and Quality (AHRQ) from June 2011 to July 2013. Due to the change in the VAP definition, the project was extended at no cost to the AHRQ for a 1-year observation period until July 2014. In this 1-year sustainability phase, four of the participating hospitals volunteered to work with the central research team to improve handoffs between ORs and ICUs. Otherwise, other project activities and supports from the central research team were largely reduced. Remaining activities included quarterly follow-up calls, continuous data collection and reporting, and provision of a communication platform via ListServe. More detail about the CSTS project is given elsewhere [7].

Data sources & outcome measures

The present study was conducted using the STS ACSD, a database that collects detailed clinical data for the purpose of assessing care and tracking performance outcomes for cardiac surgical operations performed by member hospitals. As of September 2015, there were more than 1100 hospitals or surgeon groups across 50 US states reporting data to the STS database. The database demonstrated high and increasing level of penetration at both hospital and patient level (with >90% of eligible entities participating in 2012) [24,25]. On a quarterly basis, participants receive feedback reports on major patient outcomes and care process measures in a risk-adjusted format that allows comparison with benchmarks and national standards.

We selected outcome and process measures from the STS ACSD that the CSTS interventions could potentially influence, including SSI, pneumonia among vented patients, postoperative hospital length of stay, operative mortality and Surgical Care Improvement Project (SCIP) measures. Separately, the CSTS project also collected self-reported data on SSI, CLABSI and VAP rates directly from hospitals, based on the definitions from the Centers for Disease Control and Prevention National Healthcare Safety Network and after data quality assurance training from the research team, as well as survey data on patient safety culture, teamwork and coordination from participating clinicians and staff (data not reported here). Thus, the STS ACSD analysis expanded the outcome measures for the evaluation by adding pneumonia among vented patients, postoperative hospital length of stay and operative mortality. Moreover, the definition of SSI was different between the two data sources. Furthermore, the current analysis also added process measures, including SCIP measures for compliance with appropriate selection, use and timing of prophylactic antibiotics and use of evidence-based therapies following coronary artery bypass graft (CABG) surgery, including antiplatelet, β -blocker and lipid-lowering medications at discharge. Please see [Table 1](#) for the definitions of all process and outcome measures included in the current analysis.

Propensity-score matched comparisons

Based on the intention-to-treat principle, we compared outcome measures between the 15 CSTS hospitals that initially joined the project and a group of propensity-score matched comparison hospitals. All 15 recruited CSTS hospitals were members of the STS. We identified the population of potential comparison hospitals from the STS member hospitals and surgeon groups that maintained membership from 2009 to 2013 ($n = 951$). Surgeon groups that work with multiple hospitals ($n = 8$) were excluded. All the CSTS and other hospitals were then matched to the 2009 American Hospital Association (AHA) survey database to obtain baseline hospital characteristics. Using a combination of National Provider Identifier and either exact hospital name or exact street address, 518 hospitals (54%) were successfully matched. Another 225 (23%) and 130 (13%) hospitals were manually matched using

Table 1. Definitions of outcome measures and analysis timeframe.

Measures	Operational definition	Analysis time frame
Outcome measures		
1. Surgical site infection	Deep sternal wound infection (2780/5860), mediastinitis (5870), thoracotomy infection (2790/5930), leg Infection (2800), arm infection (2801) or conduit harvest or cannulation site infection (5940)	Baseline: October 2009–September 2011 Peri-intervention: October 2011–March 2012 Post-intervention: April 2012–March 2014
2. Pneumonias among vented patients	Patient is coded as having pneumonia coded and total ventilation time is ≥ 48 h	Baseline: February 2010–January 2012 Peri-intervention: February 2012–July 2012 Postintervention: August 2012–July 2014
3. Postoperative hospital length of stay	Number of days after the procedure that the patient is in the hospital	Baseline: June 2009–May 2011 Peri-intervention: June 2011–May 2012 Post-intervention: June 2012–May 2014
4. Operative mortality	All deaths occurring during the acute episode of care in which the operation is performed even if after 30 days; or deaths occurring after discharge from the hospital, but within 30 days of the procedure	Baseline: June 2009–May 2011 Peri-intervention: June 2011–May 2012 Post-intervention: June 2012–May 2014
Process measures		
1. Surgical Care Improvement Project measures		Baseline: October 2009–September 2011 Peri-intervention: October 2011–March 2012 Postintervention: April 2012–March 2014
– INF-1	Patient receives prophylactic antibiotics within 1 h prior to surgical incision (2 h if receiving vancomycin)	
– INF-2	Patient receives prophylactic antibiotics recommended for their specific surgical procedure	
– INF-3	Patient's prophylactic antibiotics are discontinued within 24 h after surgery end time (48 h for coronary artery bypass graft surgery or other cardiac surgery)	
2. Use of evidence-based therapies		Baseline: October 2009–September 2011 Peri-intervention: October 2011–March 2012 Postintervention: April 2012–March 2014
– Antiplatelet at discharge following CABG		
– β -blocker at discharge following CABG		
– Antilipid following CABG		
CABG: Coronary artery bypass graft.		

approximate hospital name and National Provider Identifier or zip code, respectively. The above processes yielded a set of 858 STS member hospitals and surgeon groups from which we selected comparison sites.

Of the 40 variables proposed for the propensity score model, we used single imputation to address missing values in 16 variables that had <20% missing values. For nine of the variables, we found all CSTS hospitals had a single value. We then excluded from the comparison choice set all nonmatching hospitals on four important measures of the nine: community versus noncommunity hospital (community hospitals are defined as nonfederal short-term hospitals in the AHA survey database); use of electronic medical records (yes/no); provision of medical/surgical intensive care (yes/no); and provision of adult diagnostic catheterization services (yes/no). We then generated propensity scores from the remaining 31 variables to predict the probability of participation in the CSTS project, including structural characteristics (e.g., hospital bed size, teaching status), facilities (e.g., number of ICU beds), cardiac and surgical services (e.g., interventional catheterization, heart transplant), reimbursement (e.g., written contact with an health maintenance organization), staffing (e.g., total facility personnel measured by full-time equivalent) and hospital strategies (e.g., using health status indicators to design new services). The model also included median household income of the county where the hospital is located from the Area Health Resources Files to represent the socioeconomic status of the patient population the hospital served. We matched each CSTS hospital with at most four nearest neighbors within a caliper (set as one-fourth of the standard deviation of the logit of the propensity score). 12 of the 15 CSTS hospitals had four matches; one hospital had two matches and two hospitals had one match. This yielded 52 comparison hospitals.

Statistical analysis

To assess the effect of CSTS, we used a difference-in-difference analytic approach to compare changes in each outcome measure from baseline to the postintervention period between the CSTS and matched comparison hospitals. The study period was from 2009 to 2014. The evaluation time periods varied among outcome measures in order to match to the implementation of relevant intervention components. The baseline period included 2 years prior to the start of the relevant intervention. The implementation period of the intervention was 6 months for SSI, pneumonias among vented patients, SCIP measures and use of evidence-based therapies following CABG; and 1 year for postoperative length of stay and operative mortality. The postintervention period included 2 years after completion of the active intervention period (Table 1).

The main unit of analysis was the cardiac surgical procedure. Most of the outcome variables were binary. We therefore constructed logistic regression models for these measures. For postoperative hospital length of stay, we used linear regression models. Because of its skewed distribution, a log transformation was applied to normalize the data. Moreover, patients who died were coded as having the longest recorded length of stay to represent the ‘worst’ condition [26]. All models used generalized estimating equations (GEE) to account for clustering in outcomes within a hospital. Selected patient characteristics were controlled in the models, including age, gender, need for dialysis, body surface area, acuity status and previous cardiothoracic surgeries.

In addition to the intention-to-treat approach, we also used a ‘per protocol’ approach, conducting another set of analyses that added a dummy variable specifying CSTS hospitals with high versus low program adherence, which allows us to compare CSTS hospitals that adhered to the intervention protocols with other hospitals which dropped out or did not fully adopt the protocols. The data source for assessing adherence was the Team Check-up Tool [27,28], an instrument designed to survey team activities and contextual information in QI/PS projects over time. It was first administered monthly to team leaders using an online survey during CSTS implementation, and then quarterly by telephone interview during the sustainability year. One of the components of the instrument asked to what extent the unit had implemented each of the CUSP activities (i.e., no implementation, planning stage, pilot stage or fully implemented). We defined high adherence as full implementation of at least 75% of the CUSP activities reported over the entire study period across the participating units within a hospital ($n = 9$ hospitals). Intervention hospitals that dropped out ($n = 4$) or implemented less than 25% of the CUSP activities ($n = 2$) were in the low-adherence group. This second set of analyses used the same statistical models as the intention-to-treat approach with three groups of hospitals (i.e., high-adherence CSTS hospitals, low-adherence CSTS hospitals and comparison hospitals).

Results

Table 2 shows characteristics of the CSTS and matched comparison hospitals. All hospitals were community hospitals with electronic medical records; all of them provided medical/surgical intensive care and adult diagnostic catheterization services. The majority of them were large teaching hospitals with at least 500 beds. All baseline characteristics were statistically equivalent between the CSTS and matched comparison hospitals with the exception that CSTS hospitals on average had more ICU beds (187 vs 124.5 beds). Moreover, most of the study hospitals had a formal contract with a health maintenance organization and a preferred provider organization, but did not receive net patient revenue on a shared risk or capitated basis.

Table 3 presents the results based on the intention-to-treat approach, including observed values for each 6-month time period and adjusted difference-in-difference estimates. We found improving trends in some of the outcome measures among CSTS hospitals, such as reduced SSI from 0.76 per 100 cases at baseline to 0.65 2 years after program implementation and reduced pneumonia among vented patients from 1.90 per 100 vented patients to 0.76. The CSTS hospitals also showed increasing adherence to SCIP measures from 88 to 98% and a slight increase in use of evidence-based therapies, including prescription of antiplatelet medication at discharge (94–97%), and prescription of β -blocker at discharge (88–90%). However, we did not find the CSTS program had significant impact on the outcome measures over and above the results for the matched comparison group. There were generally no significant differences between changes in the CSTS hospitals and those observed in the comparison hospitals. The second set of analyses that categorized and compared high- versus low-adherence hospitals also did not suggest any differential impact on any outcome measures.

Table 2. Characteristics of Cardiovascular Surgical Translational Study hospitals and propensity score-matched comparison hospitals.

	CSTS hospitals (n = 15)	Matched comparison hospitals (n = 52)	p-value [†]
Organizational structure			
– Hospital bed size, n (%)			
– 200–299 beds	1 (6.7%)	2 (3.8%)	0.31
– 300–399 beds	1 (6.7%)	7 (13.5%)	
– 400–499 beds	1 (6.7%)	10 (19.2%)	
– 500 or more beds	12 (80.0%)	33 (63.5%)	
– Teaching hospital, n (%)	13 (86.7%)	43 (82.7%)	0.72
– Community hospital, n (%)	15 (100.0%)	52 (100.0%)	–
– Participating in a network, n (%)	7 (46.7%)	23 (44.2%)	0.87
Facilities			
– Having an electronic medical record, n (%)	15 (100.0%)	52 (100.0%)	–
– Number of ICU beds, median (IQR)	187 (82–291)	124.5 (73.5–170)	0.039
– Number of cardiac ICU beds, median (IQR)	24 (10–26)	16 (10–24)	0.43
Services			
– Number of inpatient surgical operations, median (IQR)	13,152 (7987–16,055)	9740.5 (6815.5–1118.5)	0.15
– Providing medical/surgical intensive care, n (%)	15 (100.0%)	52 (100.0%)	–
– Providing adult diagnostic catheterization, n (%)	15 (100.0%)	52 (100.0%)	–
– Providing pediatric diagnostic catheterization, n (%)	8 (53.3%)	25 (48.1%)	0.72
– Providing adult interventional catheterization, n (%)	15 (100.0%)	50 (96.2%)	0.44
– Providing pediatric interventional catheterization, n (%)	8 (53.3%)	22 (42.3%)	0.45
– Providing heart transplant, n (%)	8 (53.3%)	24 (46.2%)	0.63
– Having adult cardiac electrophysiology, n (%)	15 (100.0%)	49 (94.2%)	0.35
– Having pediatric cardiac electrophysiology, n (%)	8 (53.3%)	24 (46.2%)	0.63
– Providing burn care, n (%)	7 (46.7%)	27 (51.9%)	0.72
Reimbursement			
– Having formal written contract with an HMO, n (%)	15 (100.0%)	50 (96.2%)	0.44
– Having formal written contract with a PPO, n (%)	15 (100.0%)	50 (96.2%)	0.44
– % net patient revenue paid on a shared risk basis, n (%)	0 (0–0)	0 (0–0)	0.65
– % net patient revenue paid on a capitated basis, n (%)	0 (0–0)	0 (0–0.5)	0.72
Staffing, median (IQR)			
– Total facility personnel FTEs	6340 (4,974–8,404)	5320 (3403–7243.5)	0.34
– Physicians and dentists FTEs	17 (0–438)	9 (0–121.2)	0.74
– Registered nurses FTEs	1409 (726–2438)	1214.5 (986–1849)	0.37
– Licensed practical nurses FTEs	22 (1–45)	21.5 (8.5–56.5)	0.44
– Residents/interns FTEs	142 (0–713)	118.5 (0–547)	0.62
– Other trainees FTEs	11 (0–29)	5.5 (0–19)	0.55
Strategies			
– Working with other providers to collect, track and communicate clinical and health information across cooperating organizations, n (%)	14 (93.3%)	48 (92.3%)	0.89
– Use health status indicators to design new services or modify existing services? n (%)	15 (100.0%)	51 (98.1%)	0.59
– Having or planning to develop execute or evaluate a diversity strategy or plan, n (%)	15 (100.0%)	48 (92.3%)	0.27
Median household income of the county hospital located, median (IQR)	46,909 (40,656–56,177)	50,719 (45,632.5–59,640)	0.42

[†]From Pearson chi-square tests for categorical variables and rank tests for continuous variables

CSTS: Cardiovascular Surgical Translational Study; FTE: Full-time equivalent; HMO: Health maintenance organization; ICU: Intensive care unit; IQR: Interquartile range; PPO: Preferred provider organization.

Table 3. The impact of Cardiovascular Surgical Translational Study on outcome measures.

	n	Observed value		Adjusted difference-in-difference estimate; OR (95% CI)
		Intervention	Comparison	
Outcome measures				
Surgical site infection, per 100 cases				
– Baseline period 1	96,227	0.76	0.95	1 (Reference)
– Baseline period 2		0.78	0.84	1.18 (0.61, 2.30)
– Baseline period 3		0.87	0.54	2.10 (1.09, 4.05) [†]
– Baseline period 4		0.54	0.72	0.98 (0.36, 2.69)
– Peri-intervention		0.85	0.66	1.60 (0.71, 3.60)
– Postintervention period 1		0.31	0.77	0.49 (0.16, 1.50)
– Postintervention period 2		0.88	0.66	1.73 (0.48, 6.19)
– Postintervention period 3		0.66	0.57	1.49 (0.46, 4.81)
– Postintervention period 4		0.65	0.55	1.52 (0.60, 3.88)
Pneumonias among vented patients, per 100 cases				
– Baseline period 1	94,681	1.90	1.87	1 (Reference)
– Baseline period 2		1.60	1.96	0.86 (0.54, 1.35)
– Baseline period 3		1.65	1.63	1.03 (0.52, 2.02)
– Baseline period 4		1.54	2.01	0.76 (0.49, 1.17)
– Peri-intervention		1.58	1.64	0.95 (0.60, 1.52)
– Postintervention period 1		1.58	1.73	0.95 (0.59, 1.53)
– Postintervention period 2		1.28	1.72	0.74 (0.38, 1.45)
– Postintervention period 3		1.08	1.64	0.66 (0.34, 1.30)
– Postintervention period 4		0.76	1.35	0.60 (0.25, 1.45)
Postoperative hospital length of stay, days				
– Baseline period 1	107,418	13.2	13.5	Reference
– Baseline period 2		11.4	14.4	0.95 (0.90, 1.00)
– Baseline period 3		12.1	13.5	0.98 (0.92, 1.03)
– Baseline period 4		12.5	13.6	0.98 (0.92, 1.03)
– Peri-intervention period 1		11.9	13.8	0.92 (0.88, 0.98)
– Peri-intervention period 2		13.7	13.2	0.99 (0.93, 1.06)
– Post intervention period 1		12.6	13.0	0.98 (0.92, 1.04)
– Postintervention period 2		11.9	13.6	0.96 (0.91, 1.02)
– Postintervention period 3		12.5	12.7	0.98 (0.92, 1.04)
– Postintervention period 4		12.1	12.4	0.98 (0.90, 1.07)
Operative mortality, per 100 cases				
– Baseline period 1	107,418	2.16	2.32	1 (Reference)
– Baseline period 2		1.56	2.70	0.62 (0.41, 0.92) [†]
– Baseline period 3		1.82	2.37	0.83 (0.46, 1.49)
– Baseline period 4		1.94	2.41	0.87 (0.52, 1.46)
– Peri-intervention period 1		1.81	2.42	0.80 (0.52, 1.24)
– Peri-intervention period 2		2.45	2.18	1.14 (0.74, 1.75)
– Postintervention period 1		2.06	2.13	1.02 (0.66, 1.56)
– Postintervention period 2		1.80	2.36	0.86 (0.55, 1.32)
– Postintervention period 3		2.12	2.10	1.09 (0.72, 1.65)
– Postintervention period 4		2.00	2.00	1.17 (0.72, 1.89)
Process measures				
SCIP measures, %				
– Baseline period 1	96,227	88.1	84.5	1 (Reference)
– Baseline period 2		88.9	83.5	1.15 (0.77, 1.71)
– Baseline period 3		88.6	83.9	1.00 (0.75, 1.34)

[†]p < 0.05.

CABG: Coronary artery bypass graft; CSTS: Cardiovascular Surgical Translational Study; SCIP: Surgical Care Improvement Project.

Table 3. The impact of Cardiovascular Surgical Translational Study on outcome measures (cont.).

	n	Observed value		Adjusted difference-in-difference estimate; OR (95% CI)
		Intervention	Comparison	
– Baseline period 4		91.9	88.0	1.04 (0.72, 1.49)
– Peri-intervention		97.0	90.1	2.37 (0.35, 15.9)
– Postintervention period 1		96.5	91.7	1.56 (0.24, 10.1)
– Postintervention period 2		97.1	92.5	1.69 (0.66, 4.35)
– Postintervention period 3		98.1	92.5	2.53 (0.83, 7.71)
– Postintervention period 4		98.4	92.6	3.09 (0.74, 12.9)
Use of evidence-based therapies				
Anti-platelet at discharge following CABG, %				
– Baseline period 1	96,227	93.7	93.4	1 (Reference)
– Baseline period 2		94.7	93.8	1.18 (0.77, 1.82)
– Baseline period 3		95.1	94.3	1.14 (0.74, 1.77)
– Baseline period 4		95.6	93.6	1.44 (0.96, 2.17)
– Peri-intervention		95.0	94.4	1.13 (0.73, 1.73)
– Postintervention period 1		95.7	95.5	1.04 (0.65, 1.66)
– Postintervention period 2		96.3	95.2	1.27 (0.77, 2.10)
– Postintervention period 3		95.6	94.9	1.17 (0.75, 1.82)
– Postintervention period 4		96.9	95.7	1.36 (0.91, 2.04)
β-blocker at discharge following CABG, %				
– Baseline period 1	96,227	87.9	90.6	1 (Reference)
– Baseline period 2		87.9	90.7	1.02 (0.77, 1.37)
– Baseline period 3		86.8	90.8	0.88 (0.68, 1.14)
– Baseline period 4		85.7	91.0	0.78 (0.56, 1.10)
– Peri-intervention		88.0	90.9	1.01 (0.75, 1.36)
– Postintervention period 1		87.9	91.2	0.96 (0.70, 1.31)
– Postintervention period 2		88.5	91.0	1.04 (0.81, 1.33)
– Post intervention period 3		89.9	90.9	1.22 (0.97, 1.55)
– Postintervention period 4		90.1	91.7	1.13 (0.89, 1.44)
Antilipid following CABG, %				
– Baseline period 1		84.6	85.3	1 (Reference)
– Baseline period 2		84.7	87.0	0.90 (0.73, 1.11)
– Baseline period 3		84.1	86.2	0.91 (0.75, 1.10)
– Baseline period 4		84.3	86.7	0.87 (0.70, 1.08)
– Peri-intervention		82.3	86.1	0.79 (0.63, 0.99) [†]
– Postintervention period 1		83.8	86.4	0.89 (0.71, 1.12)
– Postintervention period 2		83.6	86.1	0.86(0.65, 1.13)
– Postintervention period 3		84.4	87.0	0.87 (0.67, 1.14)
– Postintervention period 4		84.6	86.7	0.94 (0.71, 1.24)

†p < 0.05.

CABG: Coronary artery bypass graft; CSTS: Cardiovascular Surgical Translational Study; SCIP: Surgical Care Improvement Project.

Discussion

Cardiovascular Surgical Translational Study experience

Although results showed a positive trend in SSI, pneumonia among vented patients, and Surgical Care Improvement Project (SCIP) measures among hospitals participating in a complex quality improvement intervention targeting reduction of three HAIs and improvement in the culture of patient safety, the comparative effectiveness test reported here showed no significant advantages of the intervention over an audit and passive feedback strategy. In this study, we anticipate that the intervention did not demonstrate superiority over comparisons in part due to ceiling effects in infection prevention process measures and outcomes across all sites participating in the STS. It also seems likely

that sites faced difficulty in undertaking multiple intervention strategies at the same time, potentially reducing the magnitude of effects.

This study used an external, secondary data source from a professional society to evaluate the CSTS quality improvement program in the service line of cardiac surgery. The STS ACSD permitted us to identify a concurrent comparison group for the evaluation and allowed observation over longer follow-up periods. Using the AHA survey data, we were able to conduct propensity score matching with a rich set of hospital characteristics variables and successfully identify an appropriate comparison group for a quasiexperimental design.

Moreover, the database provides opportunities to examine relevant process and outcome measures that the CSTS did not collect, but might have, if greater resources were available or were the reporting burden lighter for participating hospitals. Evaluations of QI/PS projects aimed to reduce HAIs often only assess infection rates and calculate the project's potential to influence length of stay and mortality based on estimates of illness costs from the literature. The robustness of these estimates varies, largely depending on study settings and methods for case-mix adjustment. The direct estimates with adequate case-mix adjustment from the STS ACSD provide important information that hospitals and healthcare payers may factor in when considering adopting these QI/PS initiatives.

In this particular study, the results of the STS data analysis show little evidence of definitive intervention effects over and above a comparison group. Conclusions based on these results provide additional information over the data reported directly to the research team by participating hospitals, which showed some positive effects over baseline among CSTS hospitals in CLABSI and SSI reduction at 2 years postintervention. Information on additional variables may be useful to decision makers to understand a fuller array of effects of an intervention, which may play into the decision among various QI/PS strategies. The findings of this paper suggest that participating in a national audit and feedback program with benchmarking reports may yield similar results to more intensive interventions, when they are complex and focus on multiple intervention areas. One important take away from the other assessment of the CSTS was that sites had difficulty maintaining focus on three infection types at the same time. This state of affairs exists every day for the nation's hospitals, which must address myriad safety concerns simultaneously, and which all face struggles in meeting all requirements and expectations for flawless safety consistently over time.

Several potential explanations might contribute to differences in the conclusions drawn across the two analytic approaches undertaken in the CSTS evaluation. First, we do not have information as to what extent the comparison hospitals were involved in other QI/PS efforts to improve care in cardiac surgery, and hospitals paying fees for STS reporting may likely be those committed to continuous quality improvement. The periodic audit, feedback and benchmarking on the performance measures from the STS seems likely to have motivated these hospitals to achieve high performance, possibly suggesting an above-average comparison sample. Second, the outcome measures evaluated in this study have different operational definitions and data collection protocols as compared with those in the directly reported CSTS data, and some do not reflect the exact improvement targets of the project. For example, the participating units were instructed to collect SSI rates based on the definition from the Centers for Disease Control and Prevention National Healthcare Safety Network, which is slightly different from that used in the STS ACSD in various aspects. Also, the study assessed pneumonia among vented patients, while the CSTS targeted CDC-defined VAP. Furthermore, length of stay and mortality are not direct outcomes that the CSTS aimed to improve, and in any case, it is not fully clear whether improvements in SSI, CLABSI and VAP rates may yield a measurable impact on these more distal patient outcomes.

Strengths & limitations of using an external, secondary database for QI evaluations

Based on our experiences, we note several strengths and limitations of using such databases to provide supplementary information to evaluate specific QI/PS projects as compared with using data collected directly through the project. First, it is sometimes challenging for project implementers or evaluators to identify a current comparison group due to limited resources and other providers' and organizations' unwillingness to serve in the control group. External databases, especially those from professional societies, are great sources for identifying comparison groups. Using available characteristics information for organizations, providers and/or their patients in a database, selecting an appropriate comparison group with minimal selection bias is feasible with a propensity score-matching approach. Second, these databases can usually provide high quality data with consistent variable definitions and standardized data collection protocols. They may also be more likely to provide longitudinal data with longer baseline or follow-up periods, and outcome measures that are not collected by the QI/PS projects due to resource constraints.

Third, some of these professional societies may require the QI/PS implementers to collaborate with one of their investigators or analysts; this may provide support in study design and data analysis.

On the other hand, as described in the previous paragraphs, outcome measures and their definitions from the external databases may not perfectly align with the goals of the QI/PS projects. Furthermore, comparison hospitals usually must pay a fee to join the professional societies; thus, this may affect the generalizability of the data for evaluation.

Recommendations to database owners

For owners of such databases who would like to encourage the use of the databases to help evaluate QI/PS projects, we recommend providing a one-stop website for potential users with relevant information, such as background information on the database, description of data contributors, data variable definitions and codebook, sample size, time periods of available data, data use agreement guidance, length and protocols of use application, and so forth. Moreover, information regarding data contributors' quality improvement strategies and ongoing QI/PS interventions are not usually available, but such information is critical to help users identify a comparable comparison group and/or explain the evaluation findings. Furthermore, for these databases to provide greatest utility for comparison, efforts should be made to use data definitions consistent with those in widest use, such as those of the CDC or the Society for Healthcare Epidemiology of America.

Conclusion

Clinical implications of this study are that a complex intervention to prevent three HAIs and improve safety culture was statistically equivalent to audit and feedback in cardiac surgical care. Clinical improvement can be difficult to demonstrate due to ceiling effects in infection prevention process measures and outcomes. In addition, the challenge of undertaking multiple intervention strategies at the same time can also reduce efficacy of each intervention. Clinical quality improvement initiatives are more likely to be further enhanced, sustained and disseminated when they are properly evaluated. Due to resource constraints, evaluations for QI activities often lack rigorous design, appropriate comparison groups and/or comprehensive outcome measures. Further, when the evaluators attempt to pursue comparative effectiveness evaluations with observational study design, they may face the limitation that patient characteristics and other differentials in the intervention/comparison groups can also impact outcomes, which brings the question whether the intervention itself causes the observed differences [29]. This study demonstrates a feasible and useful approach to using pre-existing, ongoing external databases, such as those of professional societies, to conduct comparative effectiveness assessment for QI/PS programs and provides valuable experiences in terms of benefits and limitations to other researchers and evaluation experts.

Summary points

- Due to resource constraints, many quality improvement (QI) evaluation designs are suboptimal.
- Using external, secondary databases to evaluate QI projects can be feasible and beneficial.
- External databases may provide a proper comparison group for QI evaluations.
- External databases may permit additional outcomes and longer follow-up periods.

Acknowledgments

The authors would like to thank the hospitals and individuals who participated in the project.

Author contributions

Conception and design: YJ Hsu, AS Kosinski, P Saha-Chaudhuri, JA Marsteller; Acquisition of the data: DE Cameron, DA Thompson, JA Marsteller; Analysis of the data: AS Kosinski, AS Wallace; Interpretation of the data: YJ Hsu, AS Kosinski, AS Wallace, P Saha-Chaudhuri, BH Chang, K Speck, MA Rosen, AP Gurses, A Xie, S Huang, DE Cameron, DA Thompson, JA Marsteller; Drafting of the article: YJ Hsu, AS Kosinski, AS Wallace, BH Chang and Revising the article critically: P Saha-Chaudhuri, K Speck, MA Rosen, AP Gurses, A Xie, S Huang, DE Cameron, DA Thompson, JA Marsteller.

Financial & competing interests disclosure

This project was supported by the Agency for Healthcare Research and Quality (grant number R18 HS19934). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

This study was approved by the Johns Hopkins University School of Medicine Institutional Review Board.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Affairs* 24(1), 138–150 (2005).
2. Grimshaw J, Eccles M, Thomas R *et al.* Toward evidence-based quality improvement. *J. Gen. Intern. Med.* 21(S2), S14–S20 (2006).
3. Shortell SM, Bennett CL, Byck GR. Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress. *Milbank Q.* 76(4), 593–624 (1998).
4. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *BMJ Qual. Saf.* 12(1), 47–52 (2003).
- **Overview of study designs for evaluating quality improvement and patient safety programs.**
5. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ* 332(7538), 413–416 (2006).
6. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q.* 82(4), 581–629 (2004).
7. Chang BH, Hsu Y-J, Rosen MA *et al.* Reducing three infections across cardiac surgery programs: a multi-site cross-unit collaboration. Unpublished Manuscript, (2018). Nov. 15.
- **Evaluation of the Cardiovascular Surgical Translational Study using data collected by the project.**
8. Berenholtz SM, Lubomski LH, Weeks K *et al.* Eliminating central line-associated bloodstream infections: a national patient safety imperative. *Infect. Control Hosp. Epidemiol.* 35(1), 56–62 (2014).
9. Pronovost P, Needham D, Berenholtz S *et al.* An intervention to decrease catheter-related bloodstream infections in the ICU. *N. Engl. J. Med.* 355(26), 2725–2732 (2006).
10. Marsteller JA, Sexton JB, Hsu YJ, Hsiao CJ. A multicenter, phased, cluster-randomized controlled trial to reduce central line-associated bloodstream infections in intensive care units. *Crit. Care Med.* 40(11), 2933–2939 (2012).
- **Randomized trial to examine the impact of the interventions in ICU setting.**
11. Horwitz LI, Krumholz HM, Green ML, Huot SJ. Transfers of patient care between house staff on internal medicine wards: a national survey. *Arch. Intern. Med.* 166(11), 1173–1177 (2006).
12. Coleman EA, Berenson RA. Lost in transition: challenges and opportunities for improving the quality of transitional care. *Ann. Intern. Med.* 141(7), 533–536 (2004).
- **Introduction to the care transition issues.**
13. Arora V, Johnson J, Lovinger D, Humphrey HJ, Meltzer DO. Communication failures in patient sign-out and suggestions for improvement: a critical incident analysis. *BMJ Qual. Saf.* 14(6), 401–407 (2005).
14. Gurses AP, Kim G, Martinez EA *et al.* Identifying and categorising patient safety hazards in cardiovascular operating rooms using an interdisciplinary approach: a multisite study. *BMJ Qual. Saf.* 21(10), 810–818 (2012).
15. Friedman DM, Berger DL. Improving team structure and communication: a key to hospital efficiency. *Arch. Surg.* 139(11), 1194–1198 (2004).
16. Petersen LA, Orav EJ, Teich JM, O'Neil AC, Brennan TA. Using a computerized sign-out program to improve continuity of inpatient care and prevent adverse events. *Jt. Comm. J. Qual. Improv.* 24(2), 77–87 (1998).
17. Van Eaton EG, Horvath KD, Lober WB, Rossini AJ, Pellegrini CA. A randomized, controlled trial evaluating the impact of a computerized rounding and sign-out system on continuity of care and resident work hours. *J. Am. Coll. Surg.* 200(4), 538–545 (2005).
18. Wahr JA, Prager RL, Abernathy JH *et al.* Patient safety in the cardiac operating room: human factors and teamwork: a scientific statement from the American Heart Association. *Circulation* 128(10), 1139–1169 (2013).
19. Werner NE, Gurses AP, Leff B, Arbaje AI. Improving care transitions across healthcare settings through a human factors approach. *J. Healthcare Qual.* 38(6), 328–343 (2016).
20. Nasarwanji MF, Badir A, Gurses AP. Standardizing handoff communication: content analysis of 27 handoff mnemonics. *J. Nurs. Care Qual.* 31(3), 238–244 (2016).

21. Hsu Y-J, Marsteller JA. Influence of the comprehensive unit-based safety program in ICUs: evidence from the Keystone ICU Project. *Am. J. Med. Qual.* (2015).
22. Pronovost P, Weast B, Rosenstein B *et al.* Implementing and validating a comprehensive unit-based safety program. *J. Patient Saf.* 1(1), 33 (2005).
23. Sawyer M, Weeks K, Goeschel CA *et al.* Using evidence, rigorous measurement, and collaboration to eliminate central catheter-associated bloodstream infections. *Crit. Care Med.* 38, S292–S298 (2010).
24. Jacobs JP, Shahian DM, Prager RL *et al.* Introduction to the STS National Database Series: Outcomes analysis, quality improvement, and patient safety. *Ann. Thorac. Surg.* 100(6), 1992–2000 (2015).
25. Jacobs JP, Shahian DM, He X *et al.* Penetration, completeness, and representativeness of The Society of Thoracic Surgeons Adult Cardiac Surgery Database. *Ann. Thorac. Surg.* 101(1), 33–41 (2016).
- **More details about the Society of Thoracic Surgeons Adult Cardiac Surgery Database.**
26. Brock GN, Barnes C, Ramirez JA, Myers J. How to handle mortality when investigating length of hospital stay and time to clinical stability. *BMC Med. Res. Methodol.* 11(1), 144 (2011).
27. Chan KS, Hsu YJ, Lubomski LH. Validity and usefulness of members reports of implementation progress in a quality improvement initiative: findings from the Team Check-up Tool (TCT). *Implement. Sci.* 6(1), 115 (2011).
28. Marsteller JA, Hsu Y-J, Chan KS, Lubomski LH. Assessing content validity and user perspectives on the Team Check-up Tool: expert survey and user focus groups. *BMJ Qual. Saf.* 26(4), 1–8 (2016).
29. Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann. Intern. Med.* 151(3), 203–205 (2009).
- **Institute of Medicine report on comparative effectiveness research.**