ORIGINAL ARTICLE

Use of Administrative Data in Efficient Auditing of Hospital-Acquired Surgical Site Infections, New York State 2009–2010

Valerie B. Haley, MS;[1] Carole Van Antwerpen, RN, BSN, CIC;[1] Boldtsetseg Tserenpuntsag, DrPH;[1] Kathleen A. Gase, MPH, CIC;[1] Peggy Hazamy, RN, BSN, CIC;[1] Diana Doughty, RN, MBA, CIC, CPHQ;[1] Marie Tsivitis, MPH, CIC;[1] Rachel L. Stricof, MPH, CIC[1,2]

OBJECTIVE. To efficiently validate the accuracy of surgical site infection (SSI) data reported to the National Healthcare Safety Network (NHSN) by New York State (NYS) hospitals.

DESIGN. Validation study.

SETTING. 176 NYS hospitals.

METHODS. NYS Department of Health staff validated the data reported to NHSN by review of a stratified sample of medical records from each hospital. The four strata were (1) SSIs reported to NHSN; (2) records with an indication of infection from diagnosis codes in administrative data but not reported to NHSN as SSIs; (3) records with discordant procedure codes in NHSN and state data sets; (4) records not in the other three strata.

RESULTS. A total of 7,059 surgical charts (6% of the procedures reported by hospitals) were reviewed. In stratum 1, 7% of reported SSIs did not meet the criteria for inclusion in NHSN and were subsequently removed. In stratum 2, 24% of records indicated missed SSIs not reported to NHSN, whereas in strata 3 and 4, only 1% of records indicated missed SSIs; these SSIs were subsequently added to NHSN. Also, in stratum 3, 75% of records were not coded for the correct NHSN procedure. Errors were highest for colon data; the NYS colon SSI rate increased by 7.5% as a result of hospital audits.

CONCLUSIONS. Audits are vital for ensuring the accuracy of hospital-acquired infection (HAI) data so that hospital HAI rates can be fairly compared. Use of administrative data increased the efficiency of identifying problems in hospitals' SSI surveillance that caused SSIs to be unreported and caused errors in denominator data.

Infect Control Hosp Epidemiol 2012;33(6):565-571

Since 2007, all hospitals in New York State (NYS) have been required by public health law to report selected hospital-acquired infections (HAIs) to the NYS Department of Health (DOH) by using the National Healthcare Safety Network (NHSN). The NHSN Patient Safety Protocol identifies standardized methods and definitions for reporting HAIs.[1] The data are used by the NYSDOH primarily to monitor trends in HAI rates and to identify hospitals with unusually high or low rates. Annual public reports summarize NYS hospital performance on HAIs.[2] It is critical that the data are reported consistently and accurately by hospitals so that comparisons among hospitals are meaningful.

The NYSDOH has been conducting on-site hospital audits since the inception of the HAI reporting program to ensure that hospitals adhere to the NHSN protocols. In 2007 and 2008, NYSDOH staff conducted reviews of a random sample of medical records. For surgical site infection (SSI) surveillance, this included colon, coronary artery bypass graft (CABG), and hip replacement/revision procedures. Given the large number of hospitals (179) in NYS and the many types of mandated HAI data, only a small sample of medical charts can be reviewed for each type of procedure in each hospital each year. Since the proportion of procedures that develop SSIs is small, the chance of identifying a missed SSI in a random sample is low. NYS sought to identify patterns in underreporting of SSIs at the time of the audit so that prompt evaluation of SSI surveillance methods could lead to improved processes and increased reporting accuracy. To increase the likelihood of identifying missed SSIs, the NYSDOH developed a method for using a secondary database to target potentially problematic records for review.

There has been controversy surrounding the use of *International Classification of Diseases, Ninth Revision* (*ICD-9*) codes from administrative data to identify HAIs.[3] The *ICD-*

Affiliations: 1. New York State Department of Health, Bureau of Healthcare Associated Infections, Albany, New York; 2. Council of State and Territorial Epidemiologists, Atlanta, Georgia.

Received November 18, 2011; accepted January 9, 2012; electronically published April 20, 2012.

 0899-823X/2012/3306-0005$15.00. DOI: 10.1086/665710

9 codes are used for reimbursement purposes and do not correspond to the HAI definitions in the NHSN surveillance protocol. However, while the codes do not define HAIs, they can be used to select records to which the formal NHSN definitions can be applied. Discharge diagnosis codes and electronic medical records are used in some hospitals to enhance surveillance of SSIs.[4-6] However, we are unaware of this technique being used as part of a continuous statewide SSI audit process. This article describes the NYSDOH 2009–2010 audit method and summarizes the impact this method had on observed SSI rates.

METHODS

Prior to each calendar quarter, NYSDOH infection control–certified HAI staff determine which hospitals will be audited during the next 3 months. The decision is based on the presence of unusually high or low SSI rates, the time elapsed since the last audit, the results of the last audit, changes in hospital staff, and problems with timeliness or completeness of reporting. The total number of charts to be reviewed per procedure is related to the number of procedures performed; that is, 9 colon charts are selected for hospitals reporting fewer than 80 colon procedures, 12 charts for those reporting 80–299 procedures, 15 charts for those reporting 300–999 procedures, and 18 charts for those reporting more than 1,000 procedures.

All NYS hospitals report administrative data on hospital inpatient stays to NYS through the Statewide Planning and Research Cooperative System (SPARCS).[7] The primary function of this system is the transmission of billing information. However, NYSDOH programs are able to access the data for approved surveillance purposes. Data on birth date, first and last names, address, gender, Social Security number, and medical record number allow the creation of a unique identifier that can be used to track hospital admissions over time. Each record contains up to 15 *ICD-9* diagnosis and procedural codes. Data sets are provided to HAI data management staff approximately every 3 months. SPARCS requires that 95% of a facility's data be submitted within 60 days after the month of patient discharge, although there are often longer delays for some facilities.

Each quarter, new and unmatched procedures reported to NHSN are matched to the updated SPARCS database. Since the majority of NHSN records are entered manually in NYS, typos prevent a perfect match of the data sets. A probabilistic algorithm adapted from Link King[8] was used to improve the match rate. First, all SPARCS records that match an NHSN record on medical record number, birth date, procedure date, first two letters of first name, or first two letters of last name within a given hospital are selected as potential matches. Then, a matching score is calculated for each potential match on the basis of a more detailed comparison of all these variables, including gender and type of procedure. If the highest score exceeds a threshold, then the record is designated a match; otherwise, it remains unmatched. Uncertain matches near the threshold are not manually reviewed. SPARCS data for the next admission by a patient to the same hospital are appended to each NHSN procedure record.

The matched NHSN-SPARCS data set is used to select records for audit in two ways. It is used as a denominator check, to identify procedures reported to NHSN that have a discordant operative procedure code listed in SPARCS, and it is used as a numerator check, to identify records with an indication of infection in SPARCS that were not reported to NHSN as SSIs.

The association between SSIs and SPARCS diagnosis codes is assessed with the likelihood-ratio-positive (LR+) statistic, which summarizes how many times more likely patients with an NHSN SSI are to have that particular SPARCS diagnosis than patients without an NHSN SSI.[9] Diagnosis codes recorded as "not present on admission" for the initial hospitalization are analyzed separately from those recorded as "present on admission" for readmissions within 30 days for colon procedures and within 1 year for CABG and hip procedures. Length of stay (LOS) for the initial admission is also evaluated. *ICD-9* codes having high likelihood ratios were used to identify procedures with potential SSIs. For the 2009 audit, a small list of diagnosis codes from the initial admission was used. Since preliminary results indicated that the method was useful in identifying reporting errors, the list of *ICD-9* codes for the initial admission was expanded, and LOS and readmission with possible infection were added. The *ICD-9* codes used each year are identified in Table 1. This updated table summarizes LR+ statistics for the cleaned 2009–2010 postaudit data and includes information on a broader set of diagnosis codes.

A stratified sample of medical records from each hospital is selected in the following order: (1) those reported to NHSN as SSIs; (2) those that were not reported to NHSN as SSIs but had a long LOS or indication of infection from diagnosis codes in administrative data; (3) those with discordant procedure codes (eg, reported to NHSN as a colon procedure but reported to SPARCS as a rectal procedure); (4) those not in the other three strata (Figure 1). Only procedures that occurred between 3 months and 1 year prior to the quarter of data preparation are selected. The minimum 3-month delay gives some time for SSIs to develop and be reported but focuses on the most recently reported data. The rolling selection process ensures that procedures are continually selected from all times of the year. This analysis includes procedures that occurred between July 2008 and December 2010.

The selected NHSN records are uploaded into an Access database. During the audit visit, the auditors can easily compare the data that were entered into NHSN with data in patient charts and can record disagreements electronically. All disagreements are reviewed with the hospital staff at the end of the visit. If the classification of cases cannot be clearly delineated by the criteria in the NHSN manual, the Centers for Disease Control and Prevention (CDC) is contacted, and

TABLE 1. Likelihood Ratio Positive (LR+) Statistics, 2009–2010

Diagnosis code, when diagnosed[a]	Used[b]	Colon LR+	CABG LR+	Hip LR+
5192: mediastinitis				
Initial	3a		31.8 (9.2, 109.7)	
Next	2a		82.7 (40.0, 170.9)	
56721: peritonitis (acute) generalized				
Initial		6.6 (4.2, 10.5)		
Next	2c	20.6 (9.8, 43.1)		
56722: peritoneal abscess				
Initial		6.7 (5.4, 8.3)		
Next	2c	8.4 (6.9, 10.3)		
5679: unspecified peritonitis				
Initial		3.1 (2.0, 5.0)		27.9 (2.9, 268.3)
Next	2c	9.3 (4.5, 19.3)		
6822: cellulitis of trunk				
Initial	1a	7.5 (6.0, 9.4)	6.6 (2.5, 17.3)	
Next	2ca	6.8 (5.5, 8.4)	29.4 (20.1, 43.1)	6.0 (1.4, 25.1)
73008: acute osteomyelitis, other sites				
Initial		7.1 (1.4, 36.7)	74.2 (19.2, 286.5)	
Next	2a	8.9 (1.6, 48.6)	116.6 (47.4, 286.8)	
73025: osteomyelitis, pelvis/thigh				
Next				22.9 (6.4, 81.7)
73028: osteomyelitis, other sites				
Initial			159.1 (18.6, 1,360)	
Next		3.1 (0.9, 10.7)	63.6 (27.3, 148.2)	
99591: sepsis without acute organ dysfunction				
Initial	1h	2.4 (1.9, 3.1)	3.4 (1.9, 5.9)	3.6 (1.3, 9.8)
Next	2ah	2.8 (2.1, 3.9)	10.3 (7.2, 14.6)	5.9 (3.8, 9.2)
99592: sepsis with acute organ dysfunction				
Initial	2ah	2.4 (2.0, 2.9)	5.0 (3.7, 6.7)	7.4 (3.9, 14.0)
Next	2a	2.1 (1.4, 3.1)	7.1 (4.7, 10.7)	5.9 (3.5, 9.9)
99666: infection due to internal joint prosthesis				
Initial				6.4 (4.3, 9.6)
Next	2h			80.0 (65.8, 97.3)
99667: infection, other internal orthopedic device				
Initial				7.6 (2.3, 24.8)
Next			10.6 (1.1, 101.8)	83.8 (50.0, 140.7)
99677: complication, internal joint prosthesis				
Initial				2.2 (1.1, 4.3)
Next	2h			18.7 (12.8, 27.4)
9974: surgical complication, digestive system				
Initial		2.3 (2.1, 2.6)	0.6 (0.3, 1.4)	1.5 (0.6, 3.6)
Next	2c	3.9 (3.1, 4.9)		1.8 (0.3, 13.2)
99811: procedure complication, hemorrhage				
Initial		0.8 (0.5, 1.3)	1.5 (1.2, 2.0)	3.6 (1.8, 6.9)
Next	2a	0.9 (0.3, 2.9)	5.7 (2.6, 12.7)	10.1 (3.0, 33.2)
99812: procedure complication, hematoma				
Initial	2h	2.6 (1.8, 3.8)	1.6 (1.1, 2.4)	12.1 (7.8, 18.8)
Next	2h	4.6 (2.7, 7.7)	10.0 (5.7, 17.4)	30.4 (21.8, 42.4)
99830: disruption of wound, unspecified				
Initial	2c	7.6 (4.2, 14.0)	12.7 (2.5, 65.5)	
Next	2a	5.9 (3.0, 11.7)	42.4 (14.8, 121.9)	251.5 (50.9, 1,244)
99831: disruption of internal surgical wound				
Initial	3ca	6.8 (4.7, 9.7)	22.3 (13.8, 35.9)	
Next	2a	7.0 (4.3, 11.4)	34.1 (23.8, 48.9)	94.3 (36.5, 243.6)
99832: disruption of external surgical wound				
Initial	3ca	7.3 (5.5, 9.6)	11.7 (7.4, 18.7)	
Next	2cah	5.5 (3.8, 8.1)	31.8 (22.3, 45.3)	127.6 (75.9, 214.5)

Table 1 (*Continued*)

Diagnosis code, when diagnosed[a]	Used[b]	Colon LR+	CABG LR+	Hip LR+
99851: infected postoperative seroma				
Initial	2c	9.8 (4.7, 20.4)	15.9 (2.9, 86.7)	41.9 (3.8, 461.6)
Next		19.3 (8.8, 42.2)	31.8 (13.3, 76.2)	119.8 (45.7, 313.6)
99859: other postoperative infection				
Initial	3cah	10.1 (9.2, 11.1)	15.0 (12.2, 18.6)	14.7 (9.1, 23.9)
Next	2cah	10.2 (9.0, 11.5)	36.6 (32.4, 41.3)	60.8 (50.1, 73.7)
99883: nonhealing surgical wound				
Initial		1.8 (0.2, 13.9)	39.8 (10.7, 147.8)	
Next	2a	4.5 (2.3, 8.6)	14.9 (8.1, 27.4)	
03811: *Staphylococcus aureus* septicemia				
Next	2a	2.0 (0.3, 15.6)	60.1 (26.9, 134.4)	30.5 (9.7, 95.5)
03840: gram-negative septicemia				
Initial	2a	2.7 (1.0, 7.8)	26.5 (8.1, 86.7)	
04109: other *Streptococcus*				
Next	2c	6.8 (3.0, 15.3)	21.2 (7.6, 59.4)	18.0 (5.2, 62.3)
04111: methicillin-susceptible *S. aureus*				
Next	2cah	10.2 (5.5, 18.8)	42.6 (29.4, 61.8)	124.9 (88.5, 176.5)
04112: methicillin-resistant *S. aureus*				
Initial		11.2 (7.1, 17.8)	6.4 (1.8, 21.9)	15.7 (4.6, 53.8)
Next	2ah	5.4 (3.6, 8.0)	23.9 (15.9, 35.9)	68.5 (52.4, 89.5)
0417: *Pseudomonas*				
Next	2a	3.5 (2.0, 6.1)	27.6 (16.4, 46.4)	20.3 (11.8, 34.9)
04185: other gram-negative organisms				
Next	2a	5.3 (3.0, 9.5)	26.0 (15.7, 43.1)	9.0 (3.9, 20.8)
LOS > 28 days				
Initial	2h	1.8 (1.4, 2.5)	2.3 (1.5, 3.5)	8.1 (3.9, 16.6)
LOS > 60 days				
Initial	2a	1.2 (0.6, 2.5)	3.7 (1.6, 8.7)	14.0 (3.1, 62.3)

NOTE. CABG, coronary artery bypass graft; LOS, length of stay. Values in parentheses are 95% confidence intervals.
[a] Initial, diagnosis recorded as "not present on admission" for the initial hospitalization; next, diagnosis recorded as "present on admission" for readmission within 30 days for colon procedures and within 1 year for CABG and hip procedures.
[b] 1, used in 2009 audit; 2, used in 2010 audit; 3, used in both 2009 and 2010 audits; a, CABG; c, colon; h, hip.

case status is clarified. Hospital staff are asked to correct all errors in the NHSN within 1 month of the on-site audit.

In this report, the determination of the auditor is considered to be the gold standard. Error rates are summarized by stratum. To determine the impact of the audit on SSI rates, we compared the 2009–2010 average SSI rate (number of SSIs per 100 procedures) after completion of the audits to what the SSI rate would have been if no audit had occurred.

RESULTS

In 2009, 160 (91%) of the 176 NYS hospitals performing colon, CABG, and hip procedures were audited. In 2010, 74% of hospitals were audited. A total of 7,059 surgical charts, representing 6% of the 112,270 procedures reported by hospitals in 2009 and 2010, were reviewed.

The completeness of the match of NHSN records to SPARCS records was dependent on the elapsed time between the date of the procedure and the date of receiving the SPARCS data. While SPARCS matches are identified for 98% of NHSN records after a year, the match rate for the time period of data selected for audit each quarter varied between 75% and 90%.

Table 1 summarizes LR+ statistics. In almost every case, *ICD-9* codes on readmission were more strongly predictive of SSI than *ICD-9* codes on initial admission. For example, patients with an NHSN CABG SSI were 32 times more likely to have mediastinitis coded on the initial SPARCS record than patients without an NHSN SSI; they were 83 times more likely to have mediastinitis coded on readmission. There was little variation in likelihood-ratio-negative (LR−) statistics (not shown). Most codes had values near 0.97; exceptions were "other postoperative infection" and "infection due to internal joint prosthesis," which had values between 0.5 and 0.8.

HAI staff reviewed each chart to determine whether the type of operative procedure entered into NHSN was actually performed. Overall, 5% of the 7,059 audited procedures were entered as NHSN procedures that were not actually performed, and the majority of these discrepancies (75%) were identified in stratum 3 (Table 2).

For procedures with correct operative procedure codes and

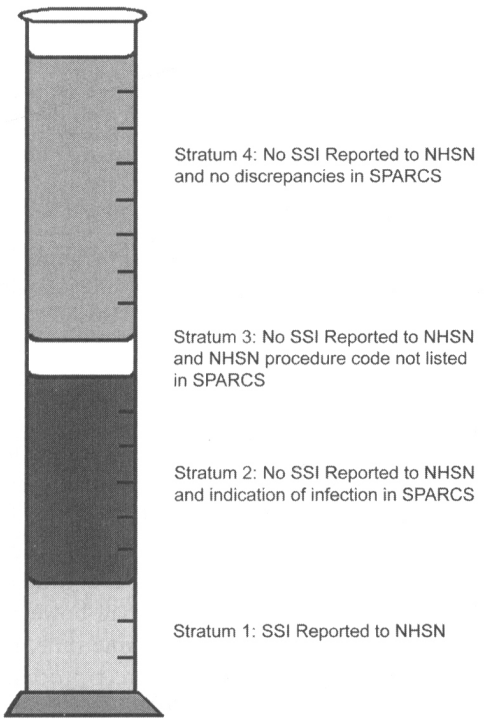

FIGURE 1. Record selection process: example of selecting 18 procedures. The process is analogous to filling a graduated cylinder, starting at the bottom with procedures already reported to NHSN as surgical site infections (SSIs), continuing with potentially problematic records, and finishing with nonproblematic records as space permits. Step 1: procedures are selected to provide a minimum ratio of 2 NHSN-reported non-SSIs per NHSN-reported SSI. If 18 procedures are needed, then 6 SSIs (stratum 1) are randomly selected. This ensures that sufficient SSIs are reviewed to validate that infection preventionists understand SSI surveillance definitions and accurately report depth of infection and microorganisms. If 6 cases are not available, the maximum number reported to NHSN are used. Step 2: the remaining number of open slots are filled with as many procedures with an indication of infection on current admission, an indication of infection on readmission, or a long hospital stay as available (stratum 2). Step 3: the remaining open slots are filled with procedures with a discrepancy in procedure code (stratum 3). Step 4: the remaining spaces are filled with randomly selected procedures that were not reported to NHSN as SSIs and either had no discrepancies in SPARCS or were not yet matched to SPARCS (stratum 4). NHSN, National Healthcare Safety Network database; SPARCS, New York's Statewide Planning and Research Cooperative System (billing database).

primary closure of the incision, HAI staff compared the patient risk factors reported to NHSN to the risk factors found in the medical charts. Error rates in entering admission date, birth date, procedure date, and gender were very low (1%–2%). Error rates for anesthesia, American Society of Anesthesiologists score, emergency, trauma, endoscope, wound class, and type of hip procedure ranged from 3% to 5%. Duration was the most inaccurate data element, with disagreements in 6% of procedures; discrepancies were most commonly due to calculation errors and reporting of the time the patient spent in the operating room rather than procedure time. In addition, reporting wound class (11% error) and endoscope use (8% error) was particularly problematic for colon procedures.

Table 3 summarizes error rates in reporting SSIs. In stratum 1, 7% of reported SSIs were found to have been entered into NHSN in error and were subsequently removed. In stratum 2, 24% of procedures had SSIs that were missed, whereas in strata 3 and 4, only 1% of procedures had missed SSIs. These SSIs were subsequently added to NHSN. The impact was greatest for colon procedures; 40% of stratum 2 colon procedure records (195 procedures) were subsequently added to NHSN as SSIs. Since the number of missed infections was greater than the number of overreported infections, the audit process resulted in increased SSI rates. The 2009–2010 colon SSI rate after completion of the audit process was 7.5% higher (95% confidence interval: 0.7%–14.8%) than it would have been if the audit had not occurred, assuming that hospitals would not have performed any additional self-validation in lieu of state audits. The impact was even greater on a per-hospital basis, since each hospital's rate is based on a smaller number of procedures, and a small change in the number of SSIs can have a large impact on the rate.

DISCUSSION

In-hospital audits are one important component of HAI data validation. The in-person visits not only serve to validate the data but also encourage open lines of communication between hospitals and NYSDOH staff, facilitate the evaluation of the surveillance methods used by hospitals to detect infections, evaluate intervention strategies, and provide on-site education related to HAI reporting and prevention. The audit process validates only data that have been reported to NHSN. Procedures that have been reported to SPARCS and not NHSN are usually not primarily closed and are therefore not reportable to NHSN. Hospitals with large percentages of missing procedures are contacted in a process separate from the audit to verify that all required procedures are reported.

Overall, 4.3% of records that were originally reported as nonevents were determined to be SSIs as a result of the audit. This percentage is higher than that in other audits that used randomly selected controls: 0.6% in Scotland,[10] 0.8% in the Netherlands,[11] and 3.3% in Australia.[12] The intense auditing process in NYS makes NYS rates appear higher than national rates because missed infections are identified and entered into the NHSN database, and training efforts increase the skills of the hospital infection preventionists (IPs), leading to better identification of HAIs. The addition or deletion of 1 or 2 infections can have a large impact on the SSI rate at an individual hospital. Therefore, it is important that a consistent audit process be applied to all hospitals every year.

This analysis showed that *ICD-9* codes from an external

TABLE 2. Accuracy of NHSN Procedure Codes by Record Selection Method

	Colon surgeries			Coronary artery bypass graft surgeries			Hip surgeries			All		
Strata	Errors	Procs	Error %	Errors	Procs	Error %	Errors	Procs	Error %	Errors	Procs	Error %
3	198	237	83.5	23	40	57.5	53	89	59.6	274	366	74.9
1, 2, 4	56	2,893	1.9	1	921	0.1	20	2,879	0.7	77	6,693	1.1
Total	254	3,130	8.1	24	961	2.5	73	2,968	2.5	351	7,059	5.0

NOTE. In stratum 3, the procedure code in NHSN does not match that in SPARCS. In strata 1, 2, and 4, the procedure code in NHSN matches that in SPARCS. Errors, number of procedures that auditor found were entered into wrong procedure category in NHSN; procs, number of procedures reviewed; NHSN, National Healthcare Safety Network database; SPARCS, New York's Statewide Planning and Research Cooperative System (billing database).

database can be used to improve reporting accuracy. LR+ statistics for many *ICD-9* codes were high; ratios above 10 are considered strong evidence to rule in SSI.[9] The *ICD-9* codes considered singly cannot be used to rule out SSI; desired tests have LR− below 0.1, and for most codes in this analysis the ratios were closer to 1.0. *ICD-9* codes on administrative data do not match the HAI definitions in the NHSN surveillance protocol; however, they can provide a quick means of identifying procedures for further review.

The secondary database increased the efficiency of error detection because more errors were detected in these records. The secondary database also increased the efficiency of the data validation process by providing an opportunity to review potentially problematic records during a regularly scheduled visit. The sample of records was more likely to contain reporting errors than a random sample, and so patterns could more easily be identified. The most common reasons for missed SSIs were misinterpretation of NHSN SSI criteria and insufficient resources dedicated to case finding. Colon data were more prone to error than hip and CABG data because additional NHSN procedures (such as small-bowel and rectal procedures) may be performed at the same time, leading to difficulty attributing the SSI to the correct procedure. Procedure entry errors occurred when IPs entered NHSN procedures on the basis of their reading of the surgeon's dictated operative report or operative log and coders assigned different *ICD-9* codes to the medical record at discharge. Approximately 25% of stratum 3 procedures were missed by medical

TABLE 3. Validation Results for Surgical Site Infection (SSI) by Type of Procedure and Stratum

	Audit results			
Procedures, strata	SSI	No SSI	Total	Interpretation
Colon procedures[a]				
Stratum 1: SSI reported to NHSN	522	64	586	10.9% of SSIs removed
Stratum 2: no SSI reported to NHSN; indication of infection in SPARCS	128	195	323	39.6% changed to SSIs
Strata 3, 4: no SSI reported to NHSN; no indication of infection in SPARCS	48	1,701	1,749	2.7% changed to SSIs
CABG procedures[b]				
Stratum 1: SSI reported to NHSN	235	4	239	1.7% of SSIs removed
Stratum 2: no SSI reported to NHSN; indication of infection in SPARCS	28	183	211	13.3% changed to SSIs
Strata 3, 4: no SSI reported to NHSN; no indication of infection in SPARCS	6	471	477	1.3% changed to SSIs
Hip procedures[c]				
Stratum 1: SSI reported to NHSN	280	10	290	3.4% of SSIs removed
Stratum 2: no SSI reported to NHSN; indication of infection in SPARCS	13	141	154	8.4% changed to SSIs
Strata 3, 4: no SSI reported to NHSN; no indication of infection in SPARCS	9	2,438	2,447	0.4% changed to SSIs
All procedures				
Stratum 1: SSI reported to NHSN	1,037	78	1,115	7.0% of SSIs removed
Stratum 2: no SSI reported to NHSN; indication of infection in SPARCS	169	519	688	24.2% changed to SSIs
Strata 3, 4: no SSI reported to NHSN; no indication of infection in SPARCS	63	4,610	4,673	1.3% changed to SSIs

NOTE. CABG, coronary artery bypass graft; SPARCS, New York's Statewide Planning and Research Cooperative System (billing database); NHSN, National Healthcare Safety Network database; NYS, New York State; CI, confidence interval.

[a] Overall NYS 2009–2010 colon SSI rate after audit (5.18%) was 7.5% (95% CI, 0.7%, 14.8%) higher than preaudit colon SSI rate (4.81%).

[b] Overall NYS 2009–2010 CABG chest site SSI rate after audit (2.29%) was 4.1% (95% CI, −7.1%, 16.8%) higher than preaudit CABG chest SSI rate (2.20%); CABG donor site SSI rate after audit (0.96%) was 3.3% (95% CI, −14.2%, 24.2%) higher than preaudit CABG donor site SSI rate (0.93%).

[c] Overall NYS 2009–2010 hip SSI rate after audit (1.18%) was 2.1% (95% CI, −8.7%, 14.2%) higher than preaudit hip SSI rate (1.16%).

record coders. Medical coders select *ICD-9* codes to maximize reimbursement rather than exactly duplicate all the diagnoses and procedures observed.

While the audit process resulted in increased SSI rates of up to 7.5%, the true SSI rate may be even higher, because a maximum of 9 months of data was evaluated at an average of 82% of hospitals each year, and the full year of follow-up after hip and CABG procedures was not completed before records were selected. There are advantages and limitations to focusing on more-recent procedures. We selected procedures that were performed 3–12 months prior to the audit to effectively evaluate current surveillance methods used by hospitals and to address timing needs for the NYSDOH Annual Report, which summarizes the most recent year of data only 9 months after the year ends. However, it is likely that additional problematic records would be identified if additional time permitted a more complete NHSN-SPARCS match and if the full year of data from all hospitals was evaluated.

The NYSDOH audit process required substantial time investment at start-up to create data management tools. Maintaining the audit program requires fewer resources, as tools exist and the process has been streamlined. NYSDOH assesses and updates the audit process each year to respond to changes in reporting requirements and the accuracy of the data. The audit process could be improved by refining the algorithm used to select potentially problematic procedures, for example, considering groups of *ICD-9* codes rather than individual codes to improve prediction and increasing the number of records with an indication of potentially missed SSIs that are reviewed. While changes in the record selection process make it more difficult to assess trends in reporting accuracy over time, they strengthen the overall accuracy of the data, so that hospital-specific HAI rates can be fairly compared.

ACKNOWLEDGMENTS

Financial support. This publication was supported in part by Cooperative Agreement 5U38HM000414-04 from the CDC.

Potential conflicts of interest. All authors report no conflicts of interest relevant to this article. All authors submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and the conflicts that the editors consider relevant to this article are disclosed here.

Address correspondence to Valerie B. Haley, MS, Corning Tower Room 2580, Albany, NY 12237 (vbh03@health.state.ny.us).

The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC.

REFERENCES

1. Centers for Disease Control and Prevention. The National Healthcare Safety Network (NHSN) manual: patient safety component protocol. http://www.cdc.gov/nhsn/TOC_PSCManual.html. Accessed September 6, 2011.
2. New York State Department of Health. Hospital-acquired infections: New York State 2010. http://www.health.state.ny.us/statistics/facilities/hospital/hospital_acquired_infections/. Published September 2011. Accessed October 12, 2011.
3. Jhung MA, Banerjee SN. Administrative coding data and health care–associated infections. *Clin Infect Dis* 2009;49(6):949–955.
4. Spolaore P, Pellizzer G, Fedeli U, et al. Linkage of microbiology reports and hospital discharge diagnoses for surveillance of surgical site infections. *J Hosp Infect* 2005;60(4):317–320.
5. Bolon MK, Hooper D, Stevenson KB, et al. Improved surveillance for surgical site infections after orthopedic implantation procedures: extending applications for automated data. *Clin Infect Dis* 2009;48(9):1223–1229.
6. Inacio MC, Paxton EW, Chen Y, et al. Leveraging electronic medical records for surveillance of surgical site infection in a total joint replacement population. *Infect Control Hosp Epidemiol* 2011;32(4):351–359.
7. New York State Department of Health. SPARCS overview. http://www.health.ny.gov/statistics/sparcs/. Accessed September 12, 2011.
8. Campbell KM. Rule your data with The Link King (a SAS/AF application for record linkage and unduplication). In: Proceedings of the 30th Annual SAS Users Group International Conference. Cary, NC: SAS Institute, 2005. Paper 020-30.
9. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329(7458):168–169.
10. McCoubrey J, Reilly J, Mullings A, Pollack KGJ, Johnston F. Validation of surgical site infection surveillance data in Scotland. *J Hosp Infect* 2005;61(3):194–200.
11. Manniën J, van der Zeeuw AE, Wille JC, van den Hof S. Validation of surgical site infection surveillance in The Netherlands. *Infect Control Hosp Epidemiol* 2007;28(1):36–41.
12. Friedman ND, Russo PL, Bull AL, Richards MJ, Kelly H. Validation of coronary artery bypass graft surgical site infection surveillance data from a statewide surveillance system in Australia. *Infect Control Hosp Epidemiol* 2007;28(7):812–817.