ORIGINAL ARTICLE

Structure, Process, and Outcome Quality of Surgical Site Infection Surveillance in Switzerland

Stefan P. Kuster, MD, MSc;[1,2,a] Marie-Christine Eisenring, CNS, ICP;[1,3,a] Hugo Sax, MD;[1,2] Nicolas Troillet, MD, MSc;[1,3,4] Swissnoso[b]

OBJECTIVE. To assess the structure and quality of surveillance activities and to validate outcome detection in the Swiss national surgical site infection (SSI) surveillance program.

DESIGN. Countrywide survey of SSI surveillance quality.

SETTING. 147 hospitals or hospital units with surgical activities in Switzerland.

METHODS. Site visits were conducted with on-site structured interviews and review of a random sample of 15 patient records per hospital: 10 from the entire data set and 5 from a subset of patients with originally reported infection. Process and structure were rated in 9 domains with a weighted overall validation score, and sensitivity, specificity, positive predictive value, and negative predictive value were calculated for the identification of SSI.

RESULTS. Of 50 possible points, the median validation score was 35.5 (range, 16.25–48.5). Public hospitals ($P < .001$), hospitals in the Italian-speaking region of Switzerland ($P = .021$), and hospitals with longer participation in the surveillance ($P = .018$) had higher scores than others. Domains that contributed most to lower scores were quality of chart review and quality of data extraction. Of 49 infections, 15 (30.6%) had been overlooked in a random sample of 1,110 patient records, accounting for a sensitivity of 69.4% (95% confidence interval [CI], 54.6%–81.7%), a specificity of 99.9% (95% CI, 99.5%–100%), a positive predictive value of 97.1% (95% CI, 85.1%–99.9%), and a negative predictive value of 98.6% (95% CI, 97.7%–99.2%).

CONCLUSIONS. Irrespective of a well-defined surveillance methodology, there is a wide variation of SSI surveillance quality. The quality of chart review and the accuracy of data collection are the main areas for improvement.

Infect Control Hosp Epidemiol 2017;1–10

Surgical site infections (SSIs) are the most common hospital-acquired infections; they are associated with increased morbidity and mortality, prolonged length of hospital stay, and increased cost.[1–6] Infection surveillance with feedback has been shown to reduce SSI rates.[7] Nationwide SSI surveillance has been performed in Switzerland since 2011.[8] In line with a broader international trend, SSI rates of each participating hospital have been made publicly available since 2014, reinforcing the need for valid data collection.

Surveillance methods should be standardized to ensure the quality and reliability of surveillance data.[9] The accuracy of the data depends on the experience, qualifications, training, and awareness of the surveillance staff.[10,11] Validation is the only independent means to determine the accuracy of surveillance data; thus, validation is essential in determining the reliability of a SSI surveillance network in which data are aggregated from multiple data collectors and are used for comparisons among hospitals.[12,13]

Validation measures are designed to detect potential sources of bias. With regard to validation of SSI surveillance, particularly selection bias (methods of patient inclusion), information and detection bias (completion of required medical information), and

Affiliations: 1. Swissnoso - National Center for Infection Control, Bern, Switzerland; 2. Division of Infectious Diseases and Hospital Epidemiology, University and University Hospital of Zurich, Zurich, Switzerland; 3. Service of Infectious Diseases, Central Institute, Valais Hospital, Sion, Switzerland; 4. Services of Infectious Diseases and Preventive Medicine, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland.

PREVIOUS PRESENTATION: These data were presented in part at the Fourth International Conference on Prevention & Infection Control (ICPIC) on June 23, 2017, in Geneva, Switzerland

[a]Authors of equal contribution.

[b]Members of Swissnoso are (in alphabetical order): Carlo Balmelli, MD, Lugano; Marie-Christine Eisenring, RN, ICP, CNS, Sion; Stephan Harbarth, MD, MS, Geneva; Stefan P. Kuster, MD, MSc, Zurich; Jonas Marschall, MD, MSc, Bern; Virginie Masserey Spicher, MD, Bern; Didier Pittet, MD, MS, Geneva; Christian Ruef, MD, Zurich; Hugo Sax, MD, Zurich; Matthias Schlegel, MD, St. Gallen; Alexander Schweiger, MD, Basel; Nicolas Troillet, MD, MSc, Sion; Andreas F. Widmer, MD, MSc, Basel; Giorgio Zanetti, MD, MSc, Lausanne.

Received March 14, 2017; accepted July 11, 2017

 DOI: 10.1017/ice.2017.169

assessment bias (correct interpretation of the study outcome) need to be considered. Although the best means of validating a SSI surveillance module is still unknown, methods of calculating sensitivity, specificity, positive predictive values, and negative predictive values (with or without structured interviews for structure and process validation) have been widely acknowledged.[14–18]

To assess the quality of the Swissnoso SSI surveillance program, structure and process for SSI surveillance were reviewed at all participating hospitals using audits and structured interviews with all persons involved in surveillance. SSI outcome data were validated by reviewing a random sample from each hospital of 10 patient records (with or without infection) and 5 additional randomly selected records of patients with infection.

MATERIALS AND METHODS

SSI Surveillance Method

In Switzerland, the first multicenter surveillance system for SSI was developed in the mid-1990s. The system was developed according to the principles of the US National Nosocomial Infections Surveillance (NNIS) system, currently known as the National Healthcare Safety Network (NHSN)[19–23] and is described in detail in a previous publication.[8] Full documentation of the surveillance methodology is available for participating hospitals on the Swissnoso website.[14]

Since 2014, starting with the 2011 data, when participation in the program became mandatory, the Swiss National Association for the Development of Quality in Hospitals and Clinics (ANQ) has openly publishing the surveillance results by hospital, including their names, NNIS/NHSN-adjusted SSI rates, and quality of surveillance as rated during onsite visits.[24]

Validation of Participating Hospitals

Since October 1, 2012, we validated the structure and process of SSI surveillance as well as SSI outcome data during dedicated validation visits for all hospitals required to participate in the program nationwide using a standardized data collection form. Hospitals were visited on site by 1 of 3 specifically trained investigators (2 registered nurses and 1 physician) with profound knowledge of the Swissnoso SSI surveillance methodology.

Surveillance Structure and Process Assessment

On-site structured interviews and observations of the surveillance process were performed with all persons involved in SSI surveillance, regardless of education, background, or the percentage of full-time equivalents ascribed for surveillance.

A weighted score was attributed according to a structured that was developed based on existing literature and expert consensus. The questionnaire covered training of persons performing the surveillance, work environment (including understaffing), potential conflicts of interest, data sources for patient selection, completeness of inclusion, completeness of required medical information (for diagnosis of SSI, during hospitalization and after discharge), quality of postdischarge surveillance, presence and type of medical supervision, and losses to follow-up (ie, attrition bias) (Table 1 and Supplemental Tables S1, S2, and S3). When different teams performed SSI surveillance at different sites or units of a hospital (eg, pediatric surgery, abdominal surgery, or cardiac surgery), each team was assessed separately and a score was attributed to each team.

SSI Outcome Validation

Patient records of electronic case report forms (eCRFs) submitted between January 1, 2009, and October 31, 2015, were eligible for review. A random sample of 10 patient records was drawn from all cases and all types of surgeries that were submitted by each respective hospital, irrespective of the presence or absence of SSI (Dataset A). In addition, for each hospital, 5 records of patients with SSI were randomly selected from all cases and all types of surgeries with originally reported infections that were submitted by the respective hospital (Dataset B).

Patient records were reviewed by the validators with assistance from on-site participants and were checked against eCRFs and paper CRFs. The outcome determination by the independent investigator was regarded as the gold standard. All cases with infection, all misclassifications (false positive and false negative), and all questionable cases needing further clarification were reviewed and resolved by consensus with 1 or 2 additional senior investigators (M.C.E. and N.T.).

Statistical Analysis

Descriptive statistics were used to outline the surveillance structures and processes of participating hospitals. Differences between groups were assessed in univariate analyses using the χ^2, Fisher's exact test, Wilcoxon rank-sum test, or Student t test, as appropriate. Multivariate linear regression analysis was used to evaluate associations between surveillance parameters (ie, language region, hospital size, hospital status [private vs public], number of hospital beds, full-time equivalents dedicated to surveillance, duration of participation in surveillance) and validation scores for the overall score and scores within individual domains. Multivariate analyses assessing the association between surveillance structure and process parameters (ie, language region, hospital size, hospital status [private vs public], number of procedures included per year, full-time equivalents dedicated to surveillance, understaffing, and overall validation score and scores of individual domains, respectively), and misclassification of infections status and types of infections, respectively, were performed using generalized estimating equations (GEE; logit link models with binomial distribution of the dependent variable and exchangeable within-group correlation structure). This method accounted for cluster effects on the hospital level, as several cases per hospital that shared the same surveillance structure and process parameters were assessed.

The quality of outcome reporting dataset comprising all randomly drawn cases (ie, cases with and without infection, as classified by the hospital) from each visited hospital yielded cases that fell into 4 categories: (1) cases reported by hospital and identified by Swissnoso validation staff as SSI cases (true positives); (2) cases not reported by hospital and ruled out as SSI cases by Swissnoso validation staff (true negatives); (3) cases reported by hospital but ruled out as SSI cases by Swissnoso validation staff (false positives); and (4) cases not reported by the hospital but identified as SSI cases by Swissnoso validation staff (false negatives). From these numbers, sensitivity, specificity, positive predictive value, and negative predictive value with 95% confidence intervals were calculated for the overall data set, with the exception of cases with incomplete information at the time of the validation visit.

All statistical analyses were performed using Stata 14.2 software (StataCorp, College Station, TX), and 2-sided P values $< .05$ were considered statistically significant.

Sample Size Estimation

In a preceding sample-size estimation, a random sample of 913 patient records was considered necessary to achieve a 95% confidence interval for sensitivity of 5% when an overall SSI prevalence of 8% and a sensitivity of 95% were assumed.[25]

RESULTS

Between October 1, 2012, and June 26, 2016, all 147 hospitals or hospital units that participated in the surveillance and had submitted cases by October 31, 2015, were visited and audited in 25 of 26 Swiss cantons. Overall, 107 hospitals (72.8%) were from the German-speaking region of Switzerland; 31 (21.1%) were from the French-speaking region, and 9 (6.1%) were from the Italian-speaking region. 96 (65.3%) were public hospitals or hospital units, and 9 (6.1%) were university affiliated. Furthermore, 87 (59.2%) hospitals participated in the surveillance for more than 3 years at the time of validation (median time of participation: 3.4 years; range, 0.8–15.8 years).

Structure and Process Validation

The characteristics of the 147 surveillance teams are shown in Table 1. The 2 surgical procedures that are most strongly represented are colon surgery (followed by 70.8% of validated hospitals) and hip prosthesis surgery (69.4%). Understaffing was noted in 34.7% of the surveillance teams, and 35.4% of medical supervisors had not undergone the required structured training in the surveillance methodology. Conflicts of interest (ie, surveillance supervised by a member of the surgical team) were detected among 11.6% of medical supervisors.

Table 2 depicts the 9 domains assessed in the process validation score, their individual scores and weights, and the unweighted mean score per domain among the 147 hospitals or hospital units. The overall mean score was 34.85 points (standard deviation [SD], 6.95 points), with a median of 35.5 points (range, 16.25–48.5 points) for a maximum of 50 points (Figure 1). The 2 domains that contributed most to lower scores were 'follow-up during hospitalization' (weighted mean difference from maximum score, 3.97 points; SD, 2.30 points) and 'data quality of eCRF compared to original data' (weighted mean difference from maximum score: 3.22 points; SD, 1.64 points).

TABLE 1. Surgical Interventions Included in the Validation and Followed by the147 Audited Surveillance Teams, and Team Characteristics, October 1, 2012, to June 26, 2016

Variable	Value
Surgical procedures followed, No. (%)	
Appendectomy	62 (42.2)
Colon surgery	104 (70.8)
Rectum surgery	15 (10.2)
Cholecystectomy	49 (33.3)
Herniorrhaphy	49 (33.3)
Gastric bypass surgery	10 (6.8)
Caesarian section	46 (31.3)
Hip prostheses	102 (69.4)
Knee prostheses	66 (44.9)
Cardiac surgery	12 (8.2)
No. of procedures included per year, median (range)	300 (15–2,000)
Full-time equivalents dedicated to surveillance, median (range)	0.2 (0.01–1.3)
Understaffing compared to volume of operations included, No. (%)	51 (34.7)
Adequate professional background of persons performing surveillance, No. (%)	134 (91.2)
Specialization of medical supervisor	
Internal medicine alone, No. (%)	32 (21.8)
Infectious diseases with or without internal medicine, No. (%)	84 (57.1)
Surgery, No. (%)	9 (6.1)
Anesthesiology, No. (%)	17 (11.6)
Other, No. (%)	5 (3.4)
Inadequate training of medical supervisor in the surveillance methodology, No. (%)	52 (35.4)
Conflict of interest of medical supervisor, No. (%)	17 (11.6)

The associations between hospital status, language region, duration of participation in the surveillance program, and hospital size with scores within individual domains are depicted in Table 3. In multivariate linear regression analysis, public hospital status ($P < .001$), Italian-speaking region ($P = .021$) and duration of participation in the surveillance programs ($P = .018$) were associated with higher validation scores, whereas hospital size was not. The number of full-time equivalents dedicated to surveillance was neither associated with the overall score nor with scores of individual domains.

SSI Outcome Validation

A total of 1,110 randomly selected clinical cases (Dataset A, ie, irrespective of the presence or absence of SSI) with complete

TABLE 2. Domains, Scores, Weights, and Mean Scores per Domain in 147 Surveillance Teams

No.	Domain	Score and Evaluation		Weight	Unweighted Mean Score (SD)	Weighted Mean Difference From Maximum Score (SD)
1	Inclusion of cases	0 points	Apparent selection bias	2	2.42 (0.78)	1.17 (1.56)
		1 point	Selection bias probable			
		2 points	Selection bias possible			
		3 points	Complete, no selection bias			
2	Medical documentation	0 points	Documentation mostly incomplete	3	1.72 (0.46)	0.83 (1.37)
		1 point	Documentation partially incomplete			
		2 points	Documentation complete			
3	Follow-up during hospitalization	0 points	No review of medical documentation	3	1.68 (0.77)	3.97 (2.30)
		1 point	Incomplete review or only in case of suspicion of infection during phone interview			
		2 points	Complete review			
		3 points	Complete review with documentation of reasoning			
4	Postdischarge surveillance, including phone interview	0 points	Not performed	1.5	1.77 (0.97)	1.85 (1.46)
		1 point	Incomplete			
		2 points	Complete, locally adapted form			
		3 points	Complete, standardized form			
5	Data quality of eCRF compared to original data	0 points	≥6 mistakes	1.5	0.85 (1.09)	3.22 (1.64)
		1 points	3–5 mistakes			
		2 points	2 mistakes			
		3 points	≤1 mistake			
6	Documentation of cases with infection	0 points	None	1.5	1.76 (0.97)	1.86 (1.46)
		1 point	Incomplete			
		2 points	Complete			
		3 points	Complete and reviewed by medical supervisor			
7	Supervision of suspected cases by medical supervisor	0 points	Never	2	2.60 (0.72)	0.81 (1.45)
		1 point	Occasionally			
		2 points	Regularly			
		3 points	Always			
8	Medical supervisor's background	0 points	None	1.5	2.50 (0.71)	0.75 (1.06)
		1 points	Surgeon or dedicated nurse alone			
		2 points	Surgeon together with internist			
		3 points	Infectious diseases specialist and/or internist			
9	Training	0 points	No participation in a training session	2.5	1.72 (0.51)	0.69 (1.29)
		1 points	Incomplete, not all staff members participating in surveillance have attended a training session			
		2 points	Complete, all staff members participating in surveillance have attended a training session			
	All domains	Maximum unweighted score: 25 points			17.02 (3.51)	15.15 (6.95)

NOTE. SD, standard deviation; eCRF, electronic case report form.

follow-up were reviewed between October 1, 2012, and June 26, 2016. The overall infection rate , as determined by the validators, was 4.4% (95% confidence interval [CI], 3.3%–5.8%). The characteristics of these cases are shown in Table 4. Overall, 15 cases (1.4%) were incorrectly classified as no infection, and 1 case (0.09%) was misclassified as an infection, accounting for a specificity of the surveillance of 99.9% (95% CI, 99.5%–100%), a sensitivity of 69.4% (95% CI, 54.6%–81.7%), a positive predictive value of 97.1% (95% CI, 85.1%–99.9%), and a negative predictive value of 98.6% (95% CI, 97.7%–99.2%). Of the 15 false negatives, 9 occurred in colon surgery cases, 3 in hip prosthesis cases, 2 in caesarean section cases, and 1 in an appendectomy case. 8 of these 15 cases (53.3%) were superficial incisional infections and 7 (46.7%) were organ-space infections. The 15 cases with missed infections were from 15 different hospitals; 4 (26.7%) were missed in private hospitals and 10 (66.7%) were missed in in non–university-affiliated public hospitals. Furthermore, 1 (6.7%) occurred in a university-affiliated hospital, corresponding to the distribution of reviewed cases among these hospital categories.

In univariate GEE, misclassification of infection status was associated with lower quality of supervision of suspected cases by a medical supervisor ($P = .009$), and unweighted mean (standard deviation) scores in cases with false-negative

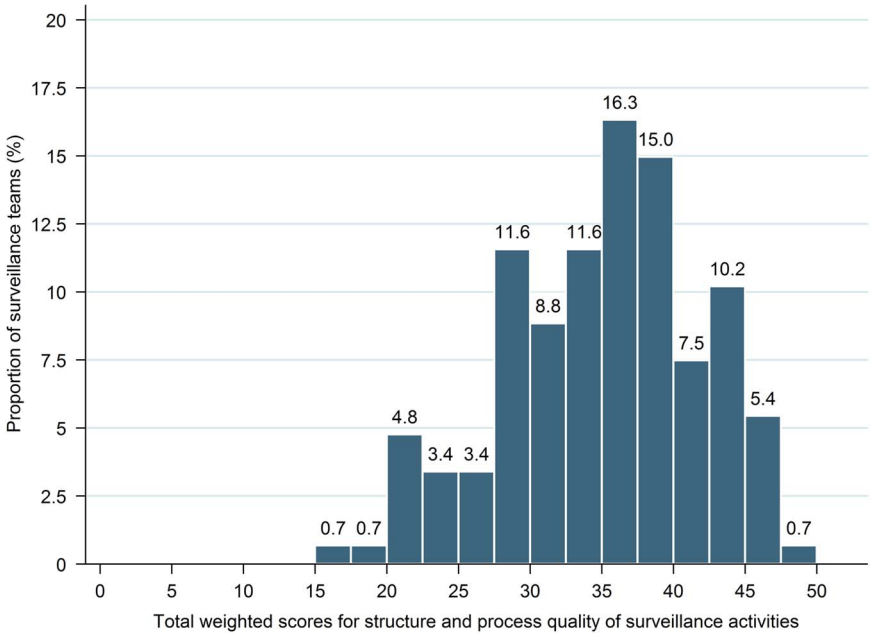

FIGURE 1. Distribution of scores in 147 participating surveillance teams audited between October 1, 2012, and June 26, 2016.

classification were 2.1 (0.93) compared to 2.61 (0.72) for cases without false-negative classification (Table 2, Domain 7). However, misclassification of infection status was not associated with the overall validation score or other domains of the score, hospital size, private hospital status, number of operations followed, medical training of the medical supervisor, or full-time equivalents dedicated for surveillance or understaffing (Supplemental Table S4).

In total, 486 cases with infections were randomly selected from 128 of 147 (87.1%) hospitals or hospital units (Dataset B, ie, randomly selected cases among those with SSI, as classified by the hospitals). The remainder of hospitals had no cases of infection at the time of the validation visit. Among these 486 cases, 204 (42.2%) were superficial incisional infections, 52 (10.8%) were deep incisional infections, and 226 (46.8%) were organ-space infections (Table 5). Misclassifications occurred in 46 (9.5%) cases. 11 superficial incisional infections were incorrectly classified as deep incisional (n = 7) or organ-space (n = 4) infections, 4 deep incisional infections were incorrectly classified as superficial incisional infections, and 31 organ-space infections were incorrectly classified as superficial incisional (n = 7) or deep incisional (n = 24) infections. 9.4% of classifications were incorrect in colon surgery, 21.2% in infections after hip arthroplasty and 18.2% in knee arthroplasty, the latter two mainly due to incorrect classification of organ-space infections as deep incisional infections. These 46 misclassifications occurred in 34 hospitals or hospital units. Of 34 hospitals, 8 (23.5%) had 2 cases with misclassification, 2 hospitals (5.9%) had 3. In univariate GEE, misclassification of types of infections was associated with lower overall validation scores ($P < .001$), higher number of operations performed ($P = .021$), lower adequacy of follow-up during hospitalization ($P = .015$), lower adequacy of documentation of cases with infection ($P < .001$), lower quality of supervision of suspected cases by the medical supervisor ($P = .003$, domain no. 7; see Table 2), lower infectious-diseases–related expertise of the medical supervisor (medical supervisor's background; $P = .007$), and lower participation in mandatory training sessions ($P = .009$). In multivariate GEE, misclassification was independently associated with lower adequacy of documentation of cases with infection ($P = .023$) (Supplemental Table S5).

DISCUSSION

Using on-site, full-day visits in all hospitals participating in SSI surveillance in Switzerland, we have demonstrated a wide variation of surveillance quality, with overall quality scores ranging from 16.25 to 48.5 (of 50) points. Room for improvement was detected for the important domains of chart review and quality of data extraction from patient charts. Overall, 15 infections were not reported, accounting for 1.4% of all cases that were classified as no SSIs by the hospitals and 30.6% of all included SSIs.

The association between Italian-speaking region and the overall score was possibly associated with the involvement of the same study personnel in the surveillance across several hospitals, guaranteeing better homogeneity in surveillance methodology. Public hospitals perform more extensive medical documentation to ensure high treatment quality across different treatment teams and thus reached higher validation scores. In private hospitals, there may be less variation in medical personnel involved in patient care; thus, thorough documentation may not be considered equally important. Last, the expertise that accumulates over the years explains the association between overall validation scores and duration of participation in the surveillance program. Our dataset showed an association between misclassifications of infection

TABLE 3. Factors Associated With Higher Scores Within Individual Domains in 147 Surveillance Teams

	Domain[a]									
Factor	Inclusion of Cases, *P* Value	Medical Documentation, *P* Value	Follow-up During Hospitalization, *P* Value	Post-Discharge Surveillance, Including Phone Interview, *P* Value	Data Quality of eCRF Compared to Original data, *P* Value	Documentation of Cases With Infection, *P* Value	Supervision of Suspected Cases by Medical Supervisor, *P* Value	Medical Supervisors Background, *P* Value	Training, *P* Value	Overall Score (All Domains), *P* Value
Univariate models										
Language region										
German	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
French	.25	.005	.26	.62	.16	.38	.42	.001	.46	.77
Italian	.05	.25	.019	.58	.42	.014	.85	.361	.28	.016
Public hospital status	.016	<.001	.55	.68	.49	.001	.037	<.001	<.001	<.001
No. of hospital beds	.004	.12	.10	.51	.36	.001	.42	.001	.07	.004
Duration of participation in surveillance	.001	.012	<.001	.96	.93	.001	.22	.011	.015	<.001
FTE dedicated for surveillance	.72	.53	.50	.15	.89	.07	.13	.86	.28	.67
Multivariate model										
Language region										
German	Ref	Ref.	Ref.	Ref.	Ref.	Ref	Ref	Ref	Ref	Ref
French	.21	**.022**	.50	.54	.53	.25	.53	**<.001**	.14	.55
Italian	.11	.162	.09	.55	.49	**.014**	1.00	.15	.19	**.021**
Public hospital status	.38	**<.001**	.71	.41	.79	**.018**	**.03**	**<.001**	**<.001**	**<.001**
No. of hospital beds	.07	.31	.33	.41	.46	**.042**	.06	.10	.87	.41
Duration of participation in surveillance	**.046**	.24	**.009**	.89	.87	.15	.27	.99	.51	**.018**

NOTE. ECRF, electronic case report form; FTE, full-time equivalents.

[a]Bold values indicate statistical significance in multivariate models.

TABLE 4. Characteristics of 1,110 Randomly Selected Clinical Cases With Complete Follow-Up Reviewed Between October 1, 2012, and June 26, 2016

Variable	Total	True Positive, No. (%)	True Negative, No. (%)	False Positive, No. (%)	False Negative, No. (%)	PPV, % (95% CI)	NPV, % (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
All surgeries	1,110	34 (3.1)	1,060 (95.4)	1 (0.1)	15 (1.4)	97.1 (85.1–99.9)	98.6 (97.7–99.2)	69.4 (54.6–81.7)	99.9 (99.5–100)
Appendectomy	110	5 (4.6)	104 (94.6)	0 (0)	1 (0.9)	100 (47.8 – 100)	99.0 (94.8–100)	83.3 (35.9–99.6)	100 (96.5–100)
Colon surgery	144	18 (12.5)	116 (80.6)	1 (0.7)	9 (6.3)	94.7 (74.0–99.9)	92.8 (86.8–96.7)	66.7 (46.0–83.5)	99.1 (95.3–100)
Rectum surgery	8	0 (0)	8 (100)	0 (0)	0 (0)	NA	100 (63.1–100)	NA	100 (63.1–100)
Cholecystectomy	123	3 (2.4)	120 (97.6)	0 (0)	0 (0)	100 (29.2–100)	100 (97.0–100)	100 (29.2–100)	100 (97.0–100)
Herniorrhaphy	165	0 (0)	165 (100)	0 (0)	0 (0)	NA	100 (97.8–100)	NA	100 (97.8–100)
Gastric bypass surgery	2	0 (0)	2 (100)	0 (0)	0 (0)	NA	100 (15.8–100)	NA	100 (15.8–100)
Caesarian section	180	2 (1.1)	176 (97.8)	0 (0)	2 (1.1)	100 (15.8–100)	98.9 (96.0–99.9)	50.0 (6.8–93.2)	100 (97.9 – 100)
Hysterectomy	11	0 (0)	11 (100)	0 (0)	0 (0)	NA	100 (71.5–100)	NA	100 (71.5–100)
Hip prosthesis	196	2 (1.0)	191 (97.5)	0 (0)	3 (1.5)	100 (15.8–100)	98.5 (95.5–99.7)	40.0 (5.3 – 85.3)	100 (98.1–100)
Knee prosthesis	130	3 (2.3)	127 (97.7)	0 (0)	0 (0)	100 (29.2–100)	100 (97.1–100)	100 (29.2–100)	100 (97.1–100)
Laminectomy	15	0 (0)	15 (100)	0 (0)	0 (0)	NA	100 (78.2–100)	NA	100 (78.2–100)
Cardiac surgery	26	1 (3.9)	25 (96.2)	0 (0)	0 (0)	100 (2.5–100)	100 (86.3–100)	100 (2.5–100)	100 (86.3–100)

NOTE. PPV, positive predictive value; NPV, negative predictive value. NA, not applicable.

TABLE 5. Characteristics of 483 Randomly Selected Cases With Originally Reported Infection at the Main Surgical Site and Complete Documentation and Follow-Up Reviewed Between October 1, 2012, and June 26, 2016

		Originally Reported Infections			Validation				
Variable	Total	Superficial Incisional, No. (%)	Deep Incisional, No. (%)	Organ-Space, No. (%)	Superficial Incisional, No. (%)	Deep Incisional, No. (%)	Organ-Space, No. (%)	False Positive, No. (%)	Misclassification of Type of Infection, No. (%)
All surgeries	483	205 (42.4)	79 (16.4)	199 (41.2)	204 (42.2)	52 (10.8)	226 (46.8)	1 (0.2)	46 (9.5)
Appendectomy	37	11 (29.7)	6 (16.2)	20 (54.1)	11 (29.7)	5 (13.5)	21 (56.8)	0 (0)	1 (2.7)
Colon surgery	213	79 (37.1)	32 (15.0)	102 (47.9)	79 (37.1)	22 (10.3)	112 (52.6)	0 (0)	20 (9.4)
Rectal surgery	5	2 (40.0)	2 (40.0)	1 (20.0)	1 (20.0)	2 (40.0)	2 (40.0)	0 (0)	2 (40.0)
Cholecystectomy	44	21 (47.7)	5 (11.4)	18 (40.9)	22 (50.0)	4 (9.1)	18 (40.9)	0 (0)	1 (2.3)
Herniorrhaphy	21	14 (66.7)	6 (28.6)	1 (4.8)	15 (71.4)	5 (23.8)	1 (4.8)	0 (0)	1 (4.8)
Gastric bypass surgery	5	2 (40.0)	1 (20.0)	2 (40.0)	2 (40.0)	1 (20.0)	2 (40.0)	0 (0)	0 (0.0)
Caesarian section	52	41 (78.9)	5 (9.6)	6 (11.5)	41 (78.9)	5 (9.6)	6 (11.5)	0 (0)	0 (0.0)
Hysterectomy	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hip prostheses	66	18 (27.3)	15 (22.7)	33 (50.0)	17 (25.8)	4 (6.1)	44 (67.7)	1 (1.5)	14 (21.2)
Knee prostheses	33	13 (39.4)	6 (18.2)	14 (42.4)	13 (39.4)	2 (6.1)	18 (54.6)	0 (0)	6 (18.2)
Laminectomiy	1	0 (0)	0 (0)	1 (100)	1	0 (0)	0 (0)	0 (0)	0 (0)
Cardiac surgery	6	4 (66.7)	1 (33.3)	1 (33.3)	3 (50.0)	2 (33.3)	1 (16.7)	0 (0)	1 (16.7)

status (ie, SSI present as compared to absent) and the quality of supervision by a medical supervisor, but not with total validation scores, hospital size, private hospital status, number of operations followed, training of the medical supervisor, full-time equivalents dedicated for surveillance or understaffing. Misclassification of the type of infection (ie, superficial incisional, deep incisional infection, or organ/space infection) was independently associated with lower adequacy of documentation.

Taken together, our findings highlight the importance of high-quality data for interfacility comparisons and, more importantly, public reporting of healthcare-associated surveillance data. Interpretive variation despite uniform surveillance definitions has been shown previously.[26,27] Furthermore, public reporting of HAI surveillance data in a system where there is great disincentive to have unfavorable outcome data may result in exclusion or reclassification of events as opposed to preventing actual negative outcomes.[28] Therefore, apart from a standardized methodology, validation of surveillance data, surveillance methods, and operations within participating facilities by an independent party are key for quality assurance under such circumstances.

As mentioned previously, the best means to validate a SSI surveillance module is still unknown. Therefore, various approaches have been proposed in the scientific literature or are available together with the surveillance methodologies, such as the validation toolkits provided by the NHSN.[29] The methods applied to validate surgical site infection surveillance (SSIS) in The Netherlands have been published in 2007 by Mannien et al.[30] Thereby, process validation by means of a structured interview as well as a prevalence study were performed. Overall positive predictive values and negative predictive values were then calculated.

Similarly, validation of SSIS data was performed in Scotland by McCoubrey et al.[12] Validation in terms of structure (ie, trained personnel and systems for SSIS, systems to ensure complete inclusion, check and confirm the number of operations) and in terms of process (ie, phone interview for identification of the systems for SSIS data collection and management at a local level) were performed. Outcome validation was conducted by calculating sensitivity, specificity, positive predictive value, and negative predictive value of the last 15 cases of SSI and 60 further randomly selected cases.

Gastmeier et al[10] compared 2 validation methods in a prevalence survey (Nosokomiale Infektionen in Deutschland Erfassung und Prävention, NIDEP) on nosocomial infections.[10] On one hand, as in other previous studies,[11,31–33] bedside validation of the 4 physician investigators was performed using 2 supervisors as the gold standard, and sensitivity and specificity were calculated. In addition, the investigators were validated using case studies.[33]

However, limitations of these approaches—including ours—are that, first, it remains unclear how and whether results of validation of structure and process by structured interviews translate to the validity of infection outcomes. Second, validation by case studies allows for the assessment of knowledge among the persons performing surveillance, but the conclusion from case study results on SSIS performance is inappropriate with regard to potential conflicts of interest because people may behave differently in the setting of case studies as compared to real-life situations in their own hospital. Third, there is no consensus about the sensitivity required to consider surveillance results to be valid. And last, given the low prevalence of SSI, large numbers of patient charts need to be reviewed to achieve an adequate level of precision.[25,34]

In conclusion, validation of process and structure of SSI surveillance and of outcome data helps identify areas for

improvement and estimate the proportion of underreporting of SSI. Validation results are reported openly together with SSI rates in Switzerland to help the public appraise the results of SSI rates in individual hospitals. However, the efforts and cost of validation are substantial; therefore, more sensitive and efficient methods for the detection of false-negative outcome measures are urgently needed. Future research should focus on the association between poor performance in process and structure measurement and reported SSI rates.

ACKNOWLEDGMENTS

We would like to thank Marylaure Dubouloz, Katja Di Salvo, and all participating hospitals for data collection and collaboration. These data were collected in collaboration with the Swiss National Association for the Development of Quality in Hospitals and Clinics (ANQ).

Financial support: No financial support was provided relevant to this article.

Potential conflicts of interest: All authors report no conflicts of interest relevant to this article.

Address correspondence to Stefan Kuster, MD, MSc, Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Raemistrasse 100 / HAL14 D6 8091 Zürich, Switzerland (stefan.kuster@usz.ch).

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2017.169.

REFERENCES

1. Kirkland KB, Briggs JP, Trivette SL, Wilkinson WE, Sexton DJ. The impact of surgical-site infections in the 1990s: attributable mortality, excess length of hospitalization, and extra costs. *Infect Control Hosp Epidemiol* 1999;20:725–730.
2. Perencevich EN, Sands KE, Cosgrove SE, Guadagnoli E, Meara E, Platt R. Health and economic impact of surgical site infections diagnosed after hospital discharge. *Emerg Infect Dis* 2003;9: 196–203.
3. Wenzel RP. The Lowbury Lecture. The economics of nosocomial infections. *J Hosp Infect* 1995;31:79–87.
4. Badia JM, Casey AL, Petrosillo N, Hudson PM, Mitchell SA, Crosby C. Impact of surgical site infection on healthcare costs and patient outcomes: a systematic review in six European countries. *J Hosp Infect* 2017;96:1–15.
5. Jenks PJ, Laurent M, McQuarry S, Watkins R. Clinical and economic burden of surgical site infection (SSI) and predicted financial consequences of elimination of SSI from an English hospital. *J Hosp Infect* 2014;86:24–33.
6. Weber WP, Zwahlen M, Reck S, et al. Economic burden of surgical site infections at a European university hospital. *Infect Control Hosp Epidemiol* 2008;29:623–629.
7. Haley RW, Culver DH, White JW, et al. The efficacy of infection surveillance and control programs in preventing nosocomial infections in US hospitals. *Am J Epidemiol* 1985;121:182–205.
8. Troillet N, Aghayev E, Eisenring MC, Widmer AF, Swissnoso. First results of the Swiss National Surgical Site Infection Surveillance Program: Who seeks shall find. *Infect Control Hosp Epidemiol* 2017;38:697–704.
9. Bruce J, Russell EM, Mollison J, Krukowski ZH. The quality of measurement of surgical wound infection as the basis for monitoring: a systematic review. *J Hosp Infect* 2001;49:99–108.
10. Gastmeier P, Kampf G, Hauer T, et al. Experience with two validation methods in a prevalence survey on nosocomial infections. *Infect Control Hosp Epidemiol* 1998;19:668–673.
11. Wenzel RP, Osterman CA, Townsend TR, et al. Development of a statewide program for surveillance and reporting of hospital-acquired infections. *J Infect Dis* 1979;140:741–746.
12. McCoubrey J, Reilly J, Mullings A, Pollock KG, Johnston F. Validation of surgical site infection surveillance data in Scotland. *J Hosp Infect* 2005;61:194–200.
13. Haley VB, Van Antwerpen C, Tserenpuntsag B, et al. Use of administrative data in efficient auditing of hospital-acquired surgical site infections, New York State 2009–2010. *Infect Control Hosp Epidemiol* 2012;33:565–571.
14. Swissnono website. https://www.swissnoso.ch/module/ssi-surveillance/material/handbuch-formulare/. Updated 2017. Accessed July 25, 2017.
15. 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;28:812–817.
16. Huotari K, Agthe N, Lyytikainen O. Validation of surgical site infection surveillance in orthopedic procedures. *Am J Infect Control* 2007;35:216–221.
17. Masia MD, Barchitta M, Liperi G, et al. Validation of intensive care unit-acquired infection surveillance in the Italian SPIN-UTI network. *J Hosp Infect* 2010;76:139–142.
18. Zuschneid I, Geffers C, Sohr D, et al. Validation of surveillance in the intensive care unit component of the German nosocomial infections surveillance system. *Infect Control Hosp Epidemiol* 2007;28:496–499.
19. Culver DH, Horan TC, Gaynes RP, et al. Surgical wound infection rates by wound class, operative procedure, and patient risk index. National Nosocomial Infections Surveillance System. *Am J Med* 1991;91:152S–157S.
20. Emori TG, Culver DH, Horan TC, et al. National nosocomial infections surveillance system (NNIS): description of surveillance methods. *Am J Infect Control* 1991;19:19–35.
21. Hubner M, Diana M, Zanetti G, Eisenring MC, Demartines N, Troillet N. Surgical site infections in colon surgery: the patient, the procedure, the hospital, and the surgeon. *Arch Surg* 2011;146:1240–1245.
22. Romy S, Eisenring MC, Bettschart V, Petignat C, Francioli P, Troillet N. Laparoscope use and surgical site infections in digestive surgery. *Ann Surg* 2008;247:627–632.
23. Staszewicz W, Eisenring MC, Bettschart V, Harbarth S, Troillet N. Thirteen years of surgical site infection surveillance in Swiss hospitals. *J Hosp Infect* 2014;88:40–47.
24. National Association for Quality Development in Hospitals and Clinics (ANQ) website. http://www.anq.ch/akutsomatik/akutsomatik-anq-hplus/. Updated 2017. Accessed July 24, 2017.
25. Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 1996;3:895–900.
26. Birgand G, Lepelletier D, Baron G, et al. Agreement among healthcare professionals in ten European countries in diagnosing case-vignettes of surgical-site infections. *PLoS One* 2013;8:e68618.
27. Nosocomial infection rates for interhospital comparison: limitations and possible solutions. A report from the National

Nosocomial Infections Surveillance (NNIS) System. *Infect Control Hosp Epidemiol* 1991;12:609–621.
28. Talbot TR, Bratzler DW, Carrico RM, et al. Public reporting of health care-associated surveillance data: recommendations from the healthcare infection control practices advisory committee. *Ann Intern Med* 2013;159:631–635.
29. National Healthcare Safety Network (NHSN) External Validation Guidance and Toolkit 2016 Centers for Disease Control and Prevention website. https://www.cdc.gov/nhsn/pdfs/validation/2016/2016-nhsn-ev-guidance.pdf. Published 2016. Accessed July 24, 2017.
30. Mannien 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;28:36–41.
31. Broderick A, Mori M, Nettleman MD, Streed SA, Wenzel RP. Nosocomial infections: validation of surveillance and computer modeling to identify patients at risk. *Am J Epidemiol* 1990;131:734–742.
32. Cardo DM, Falk PS, Mayhall CG. Validation of surgical wound surveillance. *Infect Control Hosp Epidemiol* 1993;14:211–215.
33. Larson E, Horan T, Cooper B, Kotilainen HR, Landry S, Terry B. Study of the definition of nosocomial infections (SDNI). Research Committee of the Association for Practitioners in Infection Control. *Am J Infect Control* 1991;19:259–267.
34. Carley S, Dosman S, Jones SR, Harrison M. Simple nomograms to calculate sample size in diagnostic studies. *Emerg Med J* 2005;22:180–181.