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
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.\textsuperscript{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.

\section*{Materials and Methods}

\subsection*{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)\textsuperscript{19–23} and is described in detail in a previous publication.\textsuperscript{8} Full documentation of the surveillance methodology is available for participating hospitals on the Swissnoso website.\textsuperscript{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.\textsuperscript{24}

\subsection*{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.

\subsection*{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.

\subsection*{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.).

\subsection*{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 $\chi^2$, 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.

### 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: 3.42 points; SD, 1.64 points).

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...
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 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
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 reported 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

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

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

*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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>True Positive, No. (%)</th>
<th>True Negative, No. (%)</th>
<th>False Positive, No. (%)</th>
<th>False Negative, No. (%)</th>
<th>PPV, % (95% CI)</th>
<th>NPV, % (95% CI)</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All surgeries</td>
<td>1,110</td>
<td>34 (3.1)</td>
<td>1,060 (95.4)</td>
<td>1 (0.1)</td>
<td>15 (1.4)</td>
<td>97.1 (85.1–99.9)</td>
<td>98.6 (97.7–99.2)</td>
<td>69.4 (54.6–81.7)</td>
<td>99.9 (99.5–100)</td>
</tr>
<tr>
<td>Appendectomy</td>
<td>110</td>
<td>5 (4.6)</td>
<td>104 (94.6)</td>
<td>0 (0)</td>
<td>1 (0.9)</td>
<td>100 (47.8–100)</td>
<td>99.0 (94.8–100)</td>
<td>83.3 (35.9–99.6)</td>
<td>100 (96.5–100)</td>
</tr>
<tr>
<td>Colon surgery</td>
<td>144</td>
<td>18 (12.5)</td>
<td>116 (80.6)</td>
<td>1 (0.7)</td>
<td>9 (6.3)</td>
<td>94.7 (74.0–99.9)</td>
<td>92.8 (86.8–96.7)</td>
<td>66.7 (46.0–83.5)</td>
<td>99.1 (95.3–100)</td>
</tr>
<tr>
<td>Rectum surgery</td>
<td>8</td>
<td>0 (0)</td>
<td>8 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
<td>100 (63.1–100)</td>
<td>NA</td>
<td>100 (63.1–100)</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>123</td>
<td>3 (2.4)</td>
<td>120 (97.6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>100 (29.2–100)</td>
<td>100 (97.0–100)</td>
<td>100 (29.2–100)</td>
<td>100 (97.0–100)</td>
</tr>
<tr>
<td>Herniorrhaphy</td>
<td>165</td>
<td>0 (0)</td>
<td>165 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
<td>100 (97.8–100)</td>
<td>NA</td>
<td>100 (97.8–100)</td>
</tr>
<tr>
<td>Gastric bypass surgery</td>
<td>2</td>
<td>0 (0)</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
<td>100 (15.8–100)</td>
<td>NA</td>
<td>100 (15.8–100)</td>
</tr>
<tr>
<td>Caesarian section</td>
<td>180</td>
<td>2 (1.1)</td>
<td>178 (97.8)</td>
<td>0 (0)</td>
<td>2 (1.1)</td>
<td>100 (15.8–100)</td>
<td>98.9 (96.0–99.9)</td>
<td>50.0 (6.8–93.2)</td>
<td>100 (97.9–100)</td>
</tr>
<tr>
<td>Hysterectomy</td>
<td>11</td>
<td>0 (0)</td>
<td>11 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
<td>100 (71.5–100)</td>
<td>NA</td>
<td>100 (71.5–100)</td>
</tr>
<tr>
<td>Hip prosthesis</td>
<td>196</td>
<td>2 (1.0)</td>
<td>191 (97.5)</td>
<td>0 (0)</td>
<td>3 (1.5)</td>
<td>100 (15.8–100)</td>
<td>98.5 (95.5–99.7)</td>
<td>40.0 (5.3–85.3)</td>
<td>100 (98.1–100)</td>
</tr>
<tr>
<td>Knee prosthesis</td>
<td>130</td>
<td>3 (2.3)</td>
<td>127 (97.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>100 (29.2–100)</td>
<td>100 (97.1–100)</td>
<td>100 (29.2–100)</td>
<td>100 (97.1–100)</td>
</tr>
<tr>
<td>Laminectomy</td>
<td>15</td>
<td>0 (0)</td>
<td>15 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
<td>100 (78.2–100)</td>
<td>NA</td>
<td>100 (78.2–100)</td>
</tr>
<tr>
<td>Cardiac surgery</td>
<td>26</td>
<td>1 (3.9)</td>
<td>25 (96.2)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>100 (2.5–100)</td>
<td>100 (86.3–100)</td>
<td>100 (2.5–100)</td>
<td>100 (86.3–100)</td>
</tr>
</tbody>
</table>

Note. PPV, positive predictive value; NPV, negative predictive value. NA, not applicable.
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. 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. 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. The methods applied to validate surgical site infection surveillance (SSI) in The Netherlands have been published in 2007 by Mannien et al. On one hand, as in other previous studies, the physicians were validated using case studies. 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.

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


27. Nosocomial infection rates for interhospital comparison: limitations and possible solutions. A report from the National...


