SHEA

Original Article

Semiautomated surveillance of deep surgical site infections after colorectal surgeries: A multicenter external validation of two surveillance algorithms

Janneke D.M. Verberk MSc[1,2,3] , Tjallie I.I. van der Kooi MSc[3] , David J. Hetem MD, PhD[4],

Nicolette E.W.M. Oostdam MSc[5], Mieke Noordergraaf[6], Sabine C. de Greeff PhD[3] , Marc J.M. Bonten MD, PhD[2] and

Maaike S.M. van Mourik MD, PhD[1]

[1]Department of Medical Microbiology and Infection Prevention, University Medical Centre Utrecht, Utrecht, The Netherlands, [2]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands, [3]Department of Epidemiology and Surveillance, Centre for Infectious Diseases Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands, [4]Department of Medical Microbiology and Infection Prevention, Haaglanden Medical Center, The Hague, The Netherlands, [5]Department of Infection Prevention and Control, Alrijne Zorggroep, Leiderdorp, The Netherlands and [6]Department of Medical Microbiology, Albert Schweitzer Hospital, Dordrecht, The Netherlands

Abstract

Objective: Automated surveillance methods increasingly replace or support conventional (manual) surveillance; the latter is labor intensive and vulnerable to subjective interpretation. We sought to validate 2 previously developed semiautomated surveillance algorithms to identify deep surgical site infections (SSIs) in patients undergoing colorectal surgeries in Dutch hospitals.

Design: Multicenter retrospective cohort study.

Methods: From 4 hospitals, we selected colorectal surgery patients between 2018 and 2019 based on procedure codes, and we extracted routine care data from electronic health records. Per hospital, a classification model and a regression model were applied independently to classify patients into low- or high probability of having developed deep SSI. High-probability patients need manual SSI confirmation; low-probability records are classified as no deep SSI. Sensitivity, positive predictive value (PPV), and workload reduction were calculated compared to conventional surveillance.

Results: In total, 672 colorectal surgery patients were included, of whom 28 (4.1%) developed deep SSI. Both surveillance models achieved good performance. After adaptation to clinical practice, the classification model had 100% sensitivity and PPV ranged from 11.1% to 45.8% between hospitals. The regression model had 100% sensitivity and 9.0%–14.9% PPV. With both models, <25% of records needed review to confirm SSI. The regression model requires more complex data management skills, partly due to incomplete data.

Conclusions: In this independent external validation, both surveillance models performed well. The classification model is preferred above the regression model because of source-data availability and less complex data-management requirements. The next step is implementation in infection prevention practices and workflow processes.

(Received 25 January 2022; accepted 20 May 2022; electronically published 21 June 2022)

Surgical site infections (SSIs) are among the most common healthcare-associated infections (HAIs) and result in increased costs, morbidity, postoperative length of stay, and mortality.[1–4] Reported SSI rates after colorectal surgery range from 5% to 30%, making them high-incidence procedures.[5–8] Colorectal surgeries are therefore incorporated in most SSI surveillance programs.

In most hospitals, surveillance is performed manually. However, this is experienced as labor intensive, and possibly inaccurate and is prone to subjectivity and low interrater agreement, thus limiting comparisons between hospitals.[9–11] The increasing availability of data stored in the electronic health record (EHR) offers opportunities for (partially) automating SSI surveillance, thereby reducing the workload and supporting standardization of the surveillance process. To date, several studies have published (semi)automated methods to automate SSI surveillance after colorectal surgery. Unfortunately, most of these are not feasible for Dutch hospitals (1) because they include elements that are not representative of the Dutch clinical setting and practice,

Author for correspondence: Janneke Verberk, MSc, E-mail: J.D.M.verberk-2@umcutrecht.nl. Or Maaike van Mourik, MD, PhD, E-mail: M.S.M.van.mourik-2@umcutrecht.nl

CrossMark

(2) because they have insufficient algorithm performance, (3) because processing time is delayed or (4) because they are too complex for application in real life.[12–18]

Two published semiautomated surveillance algorithms targeting deep SSI after colorectal surgery may be feasible for the Dutch setting: a classification algorithm[19] and a multivariable regression model.[20] The classification algorithm was pre-emptively designed based on clinical and surveillance practices from a French, a Spanish, and a Dutch hospital. The sensitivity was 93.3%–100% compared to manual surveillance, and the algorithm yielded a workload reduction of 73%–82%. The regression model was developed using data from a Dutch teaching hospital; we used it to predict the probability of deep SSI for each individual patient. This 5-predictor model had a sensitivity of 98.5% and a workload reduction of 63.3%.[20]

External validation or actual implementation studies of new methods for automated surveillance are scarce.[21,22] As reported by 2 systematic reviews, only 23% of the included studies used a separate validation cohort[23] and only 25% of automated surveillance were used in clinical routine.[24] Hence, knowledge about generalizability of automated surveillance models is limited, and information about the path toward actual implementation is needed.[22,25,26]

In this study, we present an independent and external validation of the previously developed classification and regression model in new cohorts of patients that underwent colorectal surgeries in different types of Dutch hospitals.[21] We investigated the feasibility of data requirements for both algorithms. If feasible and externally valid, these models can be implemented in SSI surveillance practices and workflow processes.

Methods

Study design

In this retrospective cohort study, 4 Dutch hospitals (1 academic, 2 teaching, 1 general), each with different, or different versions of, EHR systems, extracted the data needed for algorithm application. To obtain insights in hospitals' clinical practice and patient care, a questionnaire adapted from a previous study[19] was filled in by the hospital staff at the start of the study (Appendix 1 online). Feasibility of the data collection (a precondition for implementation) was evaluated by assessing the completeness of the surveillance population (denominator) and the ability of the hospitals to automatically collect case-mix variables from their EHR. Thereafter, we applied the 2 surveillance algorithms to the extracted data. Model results were compared with conventional (ie, manually annotated) surveillance.[11] Approval for this study was obtained from the institutional Review Board of the University Medical Centre Utrecht (reference no. 20-503/C) and from the local boards of directors of each participating site. Informed consent was waived given the observational and retrospective nature of this study.

Surveillance population and data collection

The hospitals identified patients aged >1 year undergoing primary colorectal resections in 2018 and/or 2019 based on procedure codes in EHR data. Hospitals could use other data sources to establish inclusion rules to construct the surveillance population and to distinguish secondary procedures or resurgeries. For the patients included in the surveillance population, structured data were extracted from the EHR including demographics, microbiological culture results, admissions (ie, prolonged length of stay or readmission), resurgeries, radiology orders, antibiotic prescriptions, and variables for case-mix correction (see Supplementary Table S1 in Appendix 2 online).

Outcome

The outcome of interest was a deep SSI (deep incisional or organ-space) within 30 days after surgery according to the Dutch surveillance protocol.[27] In short, patients having purulent drainage from the deep incision or from a drain that is placed through the wound, or having an abscess, a positive culture from the organ space, or signs and symptoms of infection in combination with wound dehiscence and a positive culture of deep soft tissue, or other evidence of infection by direct examination were considered deep SSIs. The criterion of a positive culture is not applicable in case of anastomotic leakage or perforation following the surgery. In each hospital, infection control practitioners (ICPs) manually screened patients to identify deep SSIs. This manual surveillance was considered the reference standard. All ICPs performing manual chart review received training to ensure the quality of data collection and case ascertainment.[11] Moreover, all hospitals participated in an on-site visit to validate the conventional surveillance. Details about this on-site validation visit are described below.

Feasibility of data collection

To evaluate the feasibility of the data collection, we evaluated the completeness of the surveillance population (denominator data) by comparing the patients selected by procedure codes with patients included in the reference standard. Additionally, we compared agreement between the case-mix variables (ie, risk factors: age, sex, ASA classification, wound class, stoma creation, malignancy and anastomotic leakage) that were extracted from the EHR with the case-mix variables that were collected during conventional surveillance.

Algorithm validation

Model validation of the classification model

The classification algorithm was based on the development study, using 5 elements: antibiotics, radiology orders, (re)admissions (ie, prolonged length of stay, readmissions or death), resurgeries, and microbiological cultures (Fig. 1a and Supplementary Table S2 in Appendix 2 online). All extracted data were limited to 45 days following the colorectal surgery to enable the algorithm to capture deep SSIs that developed at the end of the 30-day follow-up period. In accordance with the development study,[19] patients were classified into low probability of having had a deep SSI (≤1 element excluding microbiology, or 2–3 elements and no microbiology) and high probability of having had a deep SSI (4 elements excluding microbiology, or 2–3 elements and microbiology). High-probability patients required manual SSI confirmation, and low-probability patients were assumed free of deep SSI. If discrepancies were found between the clinical practice reported in the questionnaire and the algorithm, we evaluated whether an adaptation of the classification algorithm could have improved performance. When an algorithm element could not be computed due to incomplete data (eg, discharge date is missing so length of stay cannot be computed), the patient scored positive on that element.
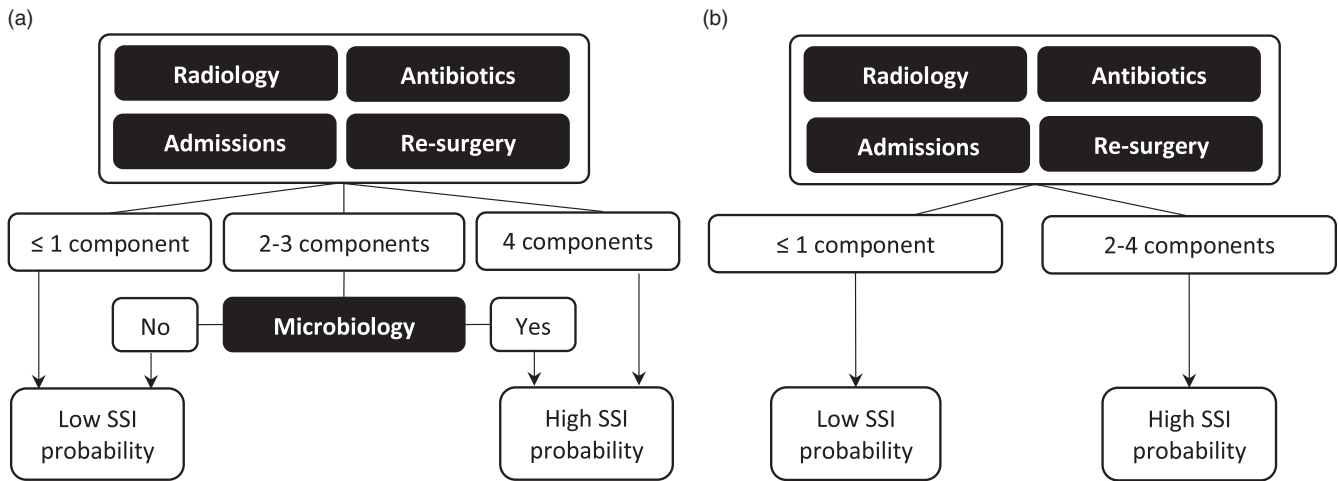
Fig. 1. Classification model. (a) Previously developed classification algorithm to classify patients with high or low probability of having had a deep surgical site infection after colorectal surgery. Figure originally published in van Rooden et al[19] and used with permission. (b) Modified classification algorithm.

$$P(DSSI) = \frac{1}{1 + e^{-LP}}$$

LP= −5.234 + 0.890 * contaminated wound (class 3) + 3.037 * re-surgery + 1.489 * readmission + 0.085 * number of postoperative days admitted to the hospital + 1.127 * mortality

Fig. 2. Previously derived prediction rule for deep surgical site infection (DSSI) after colorectal surgery. For each individual patient, the regression model returns a predicted probability of SSI which can be used to classify patients. Note. P(DSSI), probability of deep surgical site infection; LP, linear predictor.

Model validation of the regression model

The regression model utilizes wound class, hospital readmission, resurgery, postoperative length of stay, and death to calculate the probability of deep SSI. Coefficients estimated in the development setting[20] were multiplied with the predictor values of this validation cohort to estimate SSI probability (Fig. 2 and Supplementary Table S3 in Appendix 2 online). In accordance with the cutoff point in the development study, patients were classified into low probability of deep SSI (≤0.015) and high probability of deep SSI (>0.015). High-probability patients required manual SSI confirmation, whereas low-probability patients were assumed free of deep SSI. In case a predictor could not be automatically extracted by the hospital or had missing values, the predictor collected by the manual surveillance was used to evaluate algorithm performance.

On-site visit

All hospitals participated in an on-site visit to validate the conventional surveillance. This process was executed by 2 experienced surveillance advisors of the Dutch national HAI surveillance network who were blinded for the outcomes of both the reference standard and the algorithms. For each hospital, a sample of 20 patients was taken from the data according to the hierarchical rules (Fig. 3). All false-negative results were included, to confirm their deep SSI status. Additionally, records from every other group

(false-positive, true-positive, and true-negative results) were included until 20 were gathered. The group size of 20 patients was based on the time capacity of the validation team.

Statistical analyses

After data linkage, descriptive statistics were generated. To evaluate data feasibility, missing data patterns were described, and no techniques such as multiple imputation were performed to complete the data. Both models were applied to the data extractions, and results were compared with the reference standard. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and workload reduction were calculated overall and were stratified per hospital. Workload reduction was defined as the proportion of colorectal surgeries no longer requiring manual review after algorithm application. A discrepancy analysis was performed in case of any false-negative results (ie, missed deep SSI); the algorithm elements were checked in the original data. Data cleaning and statistical analyses for the classification model were carried out in SAS version 9.4 software (SAS Institute, Cary, NC). For the regression model, we used R version 3.6.1 software (R Foundation for Statistical Computing, Vienna, Austria).

Results

Feasibility of data collection

Completeness of the surveillance population

The exact surveillance population could not be reconstructed because there were no separate procedure codes or potential inclusion rules to reliably distinguish secondary procedures or resurgeries from primary procedures (range, 8.7%–22.0%, Table 1). Vice versa, 0–25% of patients in the reference standard were not identified when using inclusion rules based on procedure codes (details in Table 1). Thus, 672 colorectal surgery patients were included in this study, and 28 had deep SSIs (4.1%).

Completeness of data collection

Electronic collection of the minimum required data set from the EHR was feasible for all variables except wound class. Hospital A used text mining to establish the wound class. For hospitals B and C, wound class as collected during manual surveillance

Table 1. Overview of Data Collection and Selection of Surveillance Population

Variable	Hospital A	Hospital B	Hospital C	Hospital D
Time period extractions	2019	2018–2019[a]	2019	2019[a]
Colorectal surgeries in reference standard, no.	205	167	221	142
Colorectal surgeries extracted automatically, no.	228	159	236	148
Matched records, no.	205	124	212	131
Deep SSI in matched records, no. (%)	7 (3.4)	3 (2.4)	7 (3.3)	11 (8.3)
Records in extractions that could not be linked to reference standard, no. (%)[b]	23 (10.1)	35 (22.0)	24 (10.2)	17 (11.4)
Records in reference standard that could not be linked to extractions, no. (%)[c]	0 (0.0)	43 (25.7)	9 (4.1)	11 (7.7)

Note. SSI, surgical site infection.
[a]Until July 1, 2019.
[b]Explanation of mismatch: manual review of a random sample of these records showed these were mainly revision/secondary procedures, and for hospital C surgeries performed at another hospital location that are excluded from manual surveillance.
[c]Explanation of mismatch:
Hospital B: incorrect inclusions in reference standard as they did not meet inclusion criteria (no primary procedure)
Hospital C: These surgeries were registered as executed by internal medicine department, while for the extractions only resections performed by surgery department were selected.
Hospital D: According to the national surveillance protocol the resection with the highest risk is to be registered in case of more resections during the same surgery. Hospital included the wrong procedure in these cases.

Fig. 3. Hierarchic rules for sample selection for on-site validation of reference standard.

Hierarchic rules to select a sample from results classification model (maximum n=20):
1. All false negatives (missed SSI)
2. All superficial SSIs within the false positives
3. Maximum of 5 randomly selected false positives
4. Random 20% true positives
5. All superficial SSIs detected in the true negative group
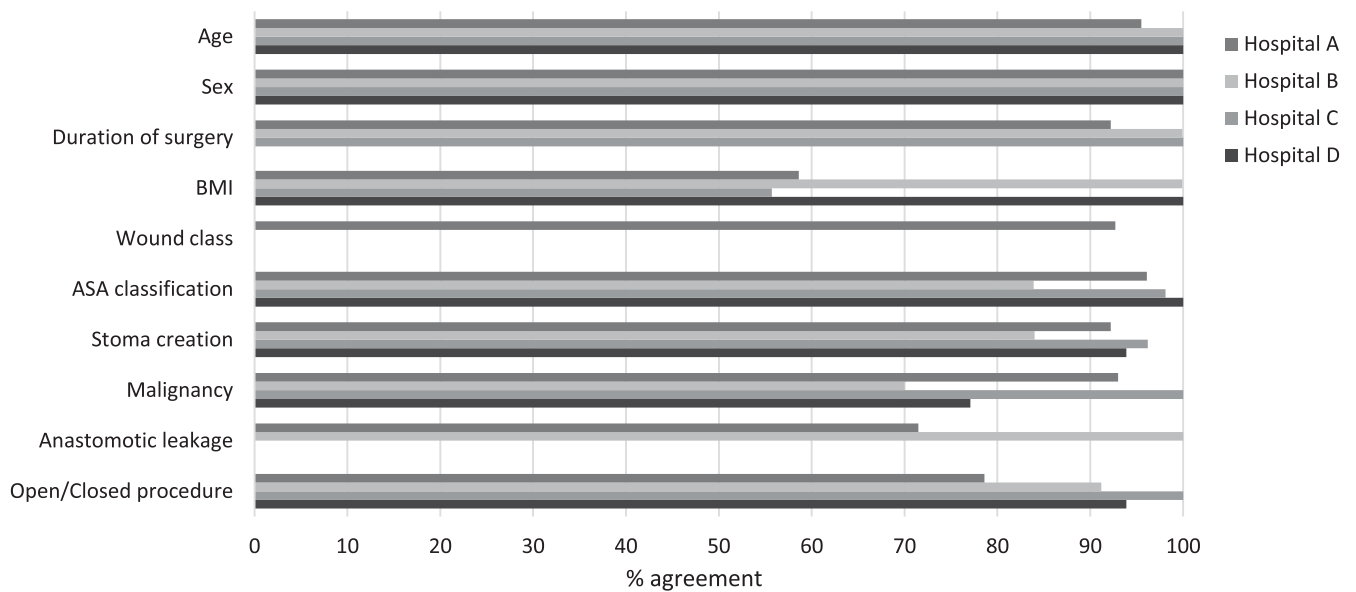6. Adding random true negatives till the number of 20 is reached

Fig. 4. Percentage agreement of risk factors extracted automatically compared to manual annotation.

(reference standard) was used. For hospital D, wound class information was not available in the source data.

Figure 4 shows the percentage of agreement between the case-mix variables extracted from the EHR and those collected manually. Disagreement was mostly related to incomplete data, either variables were not registered in the original source or were not available from source data at all.

Algorithm validation

The original classification model had an overall sensitivity of 85.7% (95% CI, 67.3%–96.0%) ranging from 72.7% to 100% between hospitals, a specificity of 92.1% (95% CI, 89.7%–94.0%), PPV of 32.0% (95% CI, 21.7%–43.8%) and an NPV of 99.3% (95% CI, 98.3%–99.8%). For the performance per hospital, see Table 2. Only

Table 2. Algorithm Performance

Variable	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	% Reduction
Classification model					
Hospital A	100 (59.0–100.0)	90.4 (85.4–94.1)	26.9 (11.6–47.8)	100 (97.9–100.0)	87.4
Hospital B	100 (29.2–100.0)	89.3 (82.3–100.0)	18.8 (4.0–45.6)	100 (96.6–100.0)	87.2
Hospital C	85.7 (42.1–99.6)	92.2 (87.6–95.5)	27.3 (10.7–50.2)	99.5 (97.1–99.9)	89.7
Hospital D	72.7 (39.0–93.9)	97.5 (92.9–99.5)	72.7 (39.0–93.9)	97.5 (92.9–99.5)	91.6
Regression model	100 (39.8–100.0)	78.1 (71.4–83.9)	9.0 (2.5–21.8)	100 (97.5–100.0)	76.5
Hospital A	100 (29.2–100.0)	78.9 (70.1–85.9)	11.1 (2.3–29.2)	100 (95.9–100.0)	76.8
Hospital B	100 (59.0–100.0)	80.0 (73.8–85.3)	14.9 (6.2–28.3)	100 (97.7–100.0)	77.3
Hospital C	NA	NA	NA	NA	NA
Hospital D	100 (39.8–100.0)	78.1 (71.4–83.9)	9.0 (2.5–21.8)	100 (97.5–100.0)	76.5
Modified classification model					
Hospital A	100 (59.0–100.0)	77.8 (71.3–83.4)	13.7 (5.7–26.3)	100 (97.6–100.0)	75.2
Hospital B	100 (29.2–100.0)	80.1 (71.9–86.9)	11.1 (2.3–29.2)	100 (96.3–100.0)	78.3
Hospital C	100 (59.0–100.0)	77.6 (71.2–83.1)	13.2 (5.4–25.3)	100 (97.7–100.0)	75.0
Hospital D	100 (71.5–100.0)	89.2 (82.2–94.1)	45.8 (25.6–67.2)	100 (96.6–100.0)	81.7

Note. CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

Table 3. Discrepancy Analyses and Explanation for Deep Surgical Site Infections (SSIs) Not Detected by Original Classification Algorithm (False-Negative Results)

False-Negative Results	Algorithm Elements, No.[a]	Missing Elements	Explanation
Patient 1, Hospital C	3	Microbiology[b] Resurgery	Treatment differed from regular treatment strategies as no reoperation was performed. Thereby, the deep SSI was scored manually based on one short clinical note stating that pus from the drain was observed.
Patient 1, Hospital D	2	Microbiology[b] Resurgery Antibiotics	No reoperation was performed. The antibiotic treatment was not identified by the algorithm as these were home-administered antibiotics, which were not included in the data selection.
Patient 2, Hospital D	3	Microbiology[b] Readmission	Reoperation took place 3 days after surgery, during the hospitalization of the index surgery; no readmission needed.
Patient 3, Hospital D	3	Microbiology[b] Resurgery	Patient had an endosponge placement; however, this reintervention is not registered as resurgery and performed as outpatient treatment by an internist, gastroenterologist or endoscopist from the gastrointestinal and liver diseases specialty while for the data extractions only resurgeries performed by same specialty as index surgery were selected

[a]Algorithm elements are radiology orders, antibiotics, (re)admissions, resurgeries, and microbiology. Patients needed 4 elements excluding microbiology, or 2-3 elements and microbiology to be classified as high probability by the algorithm. See also Fig. 1 and Appendix 2 (online).
[b]Both hospitals had reported in the questionnaires that cultures were not consistently taken in case of suspected infection.

8%–13% of the records required manual review after algorithm application. In hospitals C and D, respectively, 1 and 3 deep SSIs were missed by the algorithm (Table 3). In contrast to hospitals A and B, both hospitals had reported in the questionnaires that microbiological cultures were not consistently taken in case of suspected infection, and this was reflected in the percentage of patients meeting the microbiology element. Therefore, we adapted the algorithm and classified patients with 1 element (ie, radiology order, antibiotics, readmission, or resurgery) as low probability (Fig. 1b). This model resulted in higher sensitivity (overall sensitivity, 100%; 95% CI, 87.7%–100.0%) but at the cost of lower PPV and less workload reduction (Table 2).

The regression model could only be validated for hospitals A–C because wound class was not available for hospital D. Similar to the development study, patients with infected wounds (wound class 4) were excluded, leaving respectively 187, 116, and 207 records from hospitals A–C for analyses, including 4, 3, and 7 deep SSIs. For this model, overall sensitivity was 100% (95% CI, 76.8%–100.0%); the specificity was 76.9% (95% CI, 73.0%–80.5%); the PPV was 11.9% (95% CI, 6.6%–19.1%); and the NPV was 100% (95% CI, 99.0%–100%). With this algorithm only 22.7%–23.5% records required manual review. The results per hospital are shown in Table 2. Due to the small sample size and low number of deep SSIs, discrimination and calibration were not evaluated.

No discrepancies were found during the on-site validation visit in hospital D. In the other 3 hospitals, on-site validation revealed 5 additional deep SSIs: 2 were overlooked in the conventional surveillance and 3 were initially classified as superficial SSIs. All additional deep SSIs were classified correctly as high probability by both the (modified) classification model and the regression model. Other findings of the on-site validation of the reference standard, though not essential for the assessment of the algorithms, were reclassifications of superficial SSIs to no SSI (n = 1), missed superficial SSIs (n = 2), and incorrect inclusions (n = 8).

Discussion

This study demonstrated the external validity, both temporal and geographical, of 2 surveillance algorithms that identify patients with a high probability of deep SSI after colorectal surgery. Both had a high detection rate for deep SSI and can be used for semiautomated surveillance and, thus, to further improve efficiency and quality of SSI surveillance.

Both the classification model, especially when adapted to local practices, as well as the regression model, performed very well. To select a model for use within an organization, we considered other aspects of implementation. First, in case of incomplete data, the original development study of the regression model used multiple imputation techniques. For the classification model, the patient scored positive on the algorithm element that could not be computed due to incomplete data. This was a more convenient method for which no complex data management techniques were required. Second, according to the original study, patients with a dirty-infected wound (ie, wound class 4) were excluded from the cohort of the regression model. However, according to the national surveillance protocol, these cases should have been included in the surveillance. In addition, in 2 hospitals, wound class was not available in a structured format for automated extraction hindering algorithm application. Third, the classification model was easily be adapted to local practices. For the regression model, a sufficient sample size was required for redevelopment or recalibration in case of low predictive accuracy. This aspect may be challenging for hospitals performing few colorectal resections. Therefore, the (modified) classification model is more feasible and sustainable for real-life implementation within hospitals, improving standardization and benchmarking. We know from a previous study that the classification model has also been successful in other European countries and in low-risk surgeries such as hip and knee arthroplasties.[19,28]

For both algorithms, however, several hurdles remain for implementation. The exact surveillance population could not be automatically selected by procedure codes, but a change in the current inclusion criteria or target population could be considered. In this study, 10%–22% of surgeries detected by procedure codes did not concern a resection, were not the main indication for surgery (but performed concomitant to other intra-abdominal surgeries), or were not the first colon resection for that patient. Also, the variables necessary for case-mix adjustment are sometimes difficult to extract automatically. Although the search for a proper case-mix correction is ongoing,[14,29–32] automated extraction of a minimal set of risk factors is necessary to interpret the surveillance results and to maintain the workload reduction delivered by (semi)automated surveillance.

Two findings in this study emphasize that close monitoring, validation of algorithm components, and future maintenance are important to maintaining alignment with clinical practice and guarantee high-quality surveillance. First, as appeared from the questionnaire, 2 hospitals did not consistently obtain microbiological cultures in case of suspected deep SSI. We advise researchers to first verify whether algorithms align with clinical practice and consider adapting algorithms to differences subsequently.[23,33–35] Secondly, new treatment techniques should also be evaluated regularly and algorithms adapted accordingly. Endosponge therapy is increasingly used after anastomotic leakage; however, this intervention is often not registered or is regarded as resurgery but as outpatient treatment performed by a different specialty than the initial colorectal surgery. Each hospital should therefore periodically evaluate care practices and algorithm elements to select the appropriate resurgeries or to include recently introduced interventions, such as endosponge therapy, within the re-surgery element in the surveillance algorithm.

This study had several strengths. We performed an independent external validation in independent patient data from different types of hospitals, as well as a temporal validation. Apart from algorithm performance, automated selection of patients and case-mix variables were investigated as well, which are prerequisites for actual implementation.

This study also had several limitations. First, both algorithms targeted deep SSIs only, but in colorectal surgeries 20%–50% of SSIs are superficial.[6,36] Debate continues regarding the inclusion of superficial SSI in surveillance programs given their subjective criteria and limited clinical implications.[28,37,38] Second, we aimed to validate all published automated surveillance systems that appeared applicable to Dutch practice; however, automated surveillance systems may have been developed by commercial companies that were not published in scientific literature and were therefore not included. Third, the small sample size and low number of deep SSIs resulted in large confidence intervals for the individual hospitals and impeded the evaluation of discrimination and calibration.[39,40] Although a larger validation cohort is preferred, the numbers used in this study reflect the reality of surveillance practices. Although underpowered, the overall sensitivity and hospitals' individual point estimates were satisfying, and this study provided valuable insights into implementation. Fourth, for both manual- and semiautomated surveillance, postdischarge surveillance was limited to the initial hospital. In the Dutch setting, patients return to the operating hospital in case of complications, so this will likely not lead to underestimation of SSI rates. SSI benchmarking or widespread implementation of this semiautomated algorithm may be hampered for countries without this follow-up. Last, as actual widespread implementation of automated surveillance is still limited,[24–26] this study provides insights into validity and data requirements needed for implementation of semiautomated SSI surveillance after colorectal surgery. However, this study did not include a full feasibility study including economic,

legal, and operational assessments. We emphasize that successful implementation also depends on organizational support, information technology knowledge, staff acceptance, change management, and possibilities for integration in workflows.

In this independent external validation both approaches to semiautomated surveillance of deep SSI after colorectal surgery performed well. However, the classification model was proven preferrable to the regression model because of source data availability and less complex data-management requirements. Our results have revealed several hurdles when automating surveillance. The targeted surveillance population could not be automatically selected by procedure codes, and not all risk factors were complete or available for case-mix correction. The next step is implementation in infection prevention practices and workflow processes to automatically identify patients at increased risk of deep SSI.

References

1. Magill SS, Edwards JR, Bamberg W, *et al.* Multistate point-prevalence survey of healthcare-associated infections. *N Engl J Med* 2014;370:1198–1208.
2. Mahmoud NN, Turpin RS, Yang G, Saunders WB. Impact of surgical site infections on length of stay and costs in selected colorectal procedures. *Surg Infect* 2009;10:539–544.
3. Tanner J, Khan D, Aplin C, Ball J, Thomas M, Bankart J. Postdischarge surveillance to identify colorectal surgical site infection rates and related costs. *J Hosp Infect* 2009;72:243–250.
4. Shaw E, Gomila A, Piriz M, *et al.* Multistate modelling to estimate excess length of stay and risk of death associated with organ-space infection after elective colorectal surgery. *J Hosp Infect* 2018;100:400–405.
5. Limón E, Shaw E, Badia JM, *et al.* Postdischarge surgical site infections after uncomplicated elective colorectal surgery: impact and risk factors. The experience of the VINCat Program. *J Hosp Infect* 2014;86:127–132.
6. Referentiecijfers 2014–2018: postoperatieve wondinfecties. PREZIES website. https://www.rivm.nl/documenten/referentiecijfers-powi-2018. Published 2019. Accessed November 11, 2020.
7. Troillet N, Aghayev E, Eisenring MC, Widmer AF. First results of the Swiss National Surgical Site Infection Surveillance Program: who seeks shall find. *Infect Control Hosp Epidemiol* 2017;38:697–704.
8. Surveillance of surgical site infections in NHS hospitals in England, 2016 to 2017. Public Health England website. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/666465/SSI_annual_report_NHS_hospitals_2016-17.pdf. Published 2017. Accessed August 24, 2021.
9. Hedrick TL, Sawyer RG, Hennessy SA, Turrentine FE, Friel CM. Can we define surgical site infection accurately in colorectal surgery? *Surg Infect* 2014;15:372–376.
10. 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.
11. Verberk JDM, van Rooden SM, Hetem DJ, *et al.* Reliability and validity of multicentre surveillance of surgical site infections after colorectal surgery. *Antimicrob Resist Infect Control* 2022;11:10.
12. Cho SY, Chung DR, Choi JR, *et al.* Validation of semiautomated surgical site infection surveillance using electronic screening algorithms in 38 surgery categories. *Infect Control Hosp Epidemiol* 2018;39:931–935.
13. Gerbier-Colomban S, Bourjault M, Cetre JC, Baulieux J, Metzger MH. Evaluation study of different strategies for detecting surgical site infections using the hospital information system at Lyon University Hospital, France. *Ann Surg* 2012;255:896–900.
14. Grant R, Aupee M, Buchs NC, *et al.* Performance of surgical site infection risk prediction models in colorectal surgery: external validity assessment from three European national surveillance networks. *Infect Control Hosp Epidemiol* 2019;40:983–990.
15. Puhto T, Syrjala H. Incidence of healthcare-associated infections in a tertiary-care hospital: results from a three-year period of electronic surveillance. *J Hosp Infect* 2015;90:46–51.
16. Sohn S, Larson DW, Habermann EB, Naessens JM, Alabbad JY, Liu H. Detection of clinically important colorectal surgical site infection using Bayesian network. *J Surg Res* 2017;209:168–173.
17. Malheiro R, Rocha-Pereira N, Duro R, Pereira C, Alves CL, Correia S. Validation of a semiautomated surveillance system for surgical site infections: improving exhaustiveness, representativeness, and efficiency. *Int J Infect Dis* 2020;99:355–361.
18. van Mourik MS, van Duijn PJ, Moons KG, Bonten MJ, Lee GM. Accuracy of administrative data for surveillance of healthcare-associated infections: a systematic review. *BMJ Open* 2015;5:e008424.
19. van Rooden SM, Tacconelli E, Pujol M, *et al.* A framework to develop semiautomated surveillance of surgical site infections: an international multicenter study. *Infect Control Hosp Epidemiol* 2020;41:194–201.
20. Mulder T, Kluytmans-van den Bergh MFQ, van Mourik MSM, *et al.* A diagnostic algorithm for the surveillance of deep surgical site infections after colorectal surgery. *Infect Control Hosp Epidemiol* 2019;40:574–578.
21. Toll DB, Janssen KJM, Vergouwe Y, Moons KGM. Validation, updating, and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008;61:1085–1094.
22. Bleeker SE, Moll HA, Steyerberg EW, *et al.* External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;56:826–832.
23. Streefkerk HRA, Verkooijen RP, Bramer WM, Verbrugh HA. Electronically assisted surveillance systems of healthcare-associated infections: a systematic review. *Euro Surveill* 2020;25:1900321.
24. de Bruin JS, Seeling W, Schuh C. Data use and effectiveness in electronic surveillance of healthcare associated infections in the 21st century: a systematic review. *JAMIA* 2014;21:942–951.
25. Grota PG, Stone PW, Jordan S, Pogorzelska M, Larson E. Electronic surveillance systems in infection prevention: organizational support, program characteristics, and user satisfaction. *Am J Infect Control* 2010;38:509–514.
26. Verberk JDM, Aghdassi SJS, Abbas M, *et al.* Automated surveillance systems for healthcare-associated infections: results from a European survey and experiences from real-life utilization. *J Hosp Infect* 2022;122:35–43.
27. Case definitions SSIs. PREZIES website. https://www.rivm.nl/documenten/case-definitions-ssis. Published 2020. Accessed May 22, 2021.
28. Verberk JDM, van Rooden SM, Koek MBG, *et al.* Validation of an algorithm for semiautomated surveillance to detect deep surgical site infections after primary total hip or knee arthroplasty—a multicenter study. *Infect Control Hosp Epidemiol* 2021;42:69–74.
29. Young H, Knepper B, Moore EE, Johnson JL, Mehler P, Price CS. Surgical site infection after colon surgery: National Healthcare Safety Network risk factors and modeled rates compared with published risk factors and rates. *J Am Coll Surg* 2012;214:852–859.
30. Caroff DA, Wang R, Zhang Z, *et al.* The limited utility of ranking hospitals based on their colon surgery infection rates. *Clin Infect Dis* 2021;72:90–98.
31. Wick EC, Vogel JD, Church JM, Remzi F, Fazio VW. Surgical site infections in a "high outlier" institution: are colorectal surgeons to blame? *Dis Colon Rectum* 2009;52:374–379.

32. Watanabe M, Suzuki H, Nomura S, *et al.* Performance assessment of the risk index category for surgical site infection after colorectal surgery. *Surg Infect* 2015;16:84–89.

33. Leal J, Laupland KB. Validity of electronic surveillance systems: a systematic review. *J Hosp Infect* 2008;69:220–229.

34. Gastmeier P, Bräuer H, Hauer T, Schumacher M, Daschner F, Rüden H. How many nosocomial infections are missed if identification is restricted to patients with either microbiology reports or antibiotic administration? *Infect Control Hosp Epidemiol* 1999;20:124–127.

35. Freeman R, Moore LS, Garcia Alvarez L, Charlett A, Holmes A. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. *J Hosp Infect* 2013;84:106–119.

36. *Healthcare-Associated Infections: Surgical Site Infections—Annual Epidemiological Report for 2017*. Stockholm: European Centre for Disease Prevention and Control; 2019.

37. Skube SJ, Hu Z, Arsoniadis EG, *et al.* Characterizing surgical site infection signals in clinical notes. *Stud Health Technol Informat* 2017;245:955–959.

38. Kao LS, Ghaferi AA, Ko CY, Dimick JB. Reliability of superficial surgical site infections as a hospital quality measure. *J Am Coll Surg* 2011;213:231–235.

39. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles' heel of predictive analytics. *BMC Med* 2019;17:230.

40. Riley RD, Debray TPA, Collins GS, *et al.* Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Statist Med* 2021;40:4230–4251.