

Original Article

A framework to develop semiautomated surveillance of surgical site infections: An international multicenter study

Stephanie M. van Rooden PhD¹ , Evelina Tacconelli MD, PhD^{2,3}, Miquel Pujol MD, PhD⁴, Aina Gomila MD, PhD^{4,5}, Jan A. J. W. Kluytmans MD, PhD^{1,6,7}, Jannie Romme⁶, Gonny Moen⁶, Elodie Couvé-Deacon MD, PhD⁸, Camille Bataille⁸, Jesús Rodríguez Baño MD, PhD^{9,10}, Joaquín Lanz MD^{9,10} and Maaïke S.M. van Mourik MD, PhD¹¹

¹Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, The Netherlands, ²Division of Infectious Disease, Department of Internal Medicine I, Tübingen University Hospital, Tübingen, Germany, ³Division of Infectious Diseases, Department of Diagnostic and Public Health, University of Verona, Verona, Italy, ⁴Infectious Diseases Department, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Bellvitge University Hospital, Barcelona, Spain, ⁵Department of Infectious Diseases, Corporació Sanitària Parc Taulí, Sabadell Barcelona, Spain, ⁶Department of Microbiology and Infection Control, Amphia Hospital, Breda, The Netherlands, ⁷Microvida Laboratory for Microbiology, Amphia Hospital, Breda, The Netherlands, ⁸CHU Limoges Laboratory of Bacteriology-Virology-Hygiene, F-87000 Limoges, France, ⁹Unidad Clínica de Enfermedades Infecciosas, Microbiología y Medicina Preventiva, Hospital Universitario Virgen Macarena, ¹⁰Department of Medicine, Universidad de Sevilla/Instituto de Biomedicina de Sevilla, Seville, Spain and ¹¹Department of Medical Microbiology and Infection Control, University Medical Center Utrecht, University of Utrecht, Utrecht, The Netherlands

Abstract

Objective: Automated surveillance of healthcare-associated infections reduces workload and improves standardization, but it has not yet been adopted widely. In this study, we assessed the performance and feasibility of an easy implementable framework to develop algorithms for semiautomated surveillance of deep incisional and organ-space surgical site infections (SSIs) after orthopedic, cardiac, and colon surgeries.

Design: Retrospective cohort study in multiple countries.

Methods: European hospitals were recruited and selected based on the availability of manual SSI surveillance data from 2012 onward (reference standard) and on the ability to extract relevant data from electronic health records. A questionnaire on local manual surveillance and clinical practices was administered to participating hospitals, and the information collected was used to pre-emptively design semiautomated surveillance algorithms standardized for multiple hospitals and for center-specific application. Algorithm sensitivity, positive predictive value, and reduction of manual charts requiring review were calculated. Reasons for misclassification were explored using discrepancy analyses.

Results: The study included 3 hospitals, in the Netherlands, France, and Spain. Classification algorithms were developed to indicate procedures with a high probability of SSI. Components concerned microbiology, prolonged length of stay or readmission, and reinterventions. Antibiotics and radiology ordering were optional. In total, 4,770 orthopedic procedures, 5,047 cardiac procedures, and 3,906 colon procedures were analyzed. Across hospitals, standardized algorithm sensitivity ranged between 82% and 100% for orthopedic surgery, between 67% and 100% for cardiac surgery, and between 84% and 100% for colon surgery, with 72%–98% workload reduction. Center-specific algorithms had lower sensitivity.

Conclusions: Using this framework, algorithms for semiautomated surveillance of SSI can be successfully developed. The high performance of standardized algorithms holds promise for large-scale standardization.

(Received 16 August 2019; accepted 25 October 2019; electronically published 30 December 2019)

The burden of healthcare-associated infections (HAIs) is considerable. In European hospitals, ~6% of patients hospitalized are affected by an HAI, and surgical site infections (SSIs) are among the most

common complications.^{1,2} Feedback of infection rates from surveillance data to clinicians and other stakeholders is a cornerstone of HAI prevention programs, and participation in surveillance programs contributes to the reduction of HAI incidence.^{3–5}

Conventional retrospective surveillance relies on manual chart review to identify HAIs; however, this process is error prone and time-consuming.^{6,7} With the large-scale adoption of electronic health records (EHRs), automation of the surveillance process is increasingly feasible. Semiautomated surveillance is commonly used: routine care data stored in the EHR serve as input for an algorithm that indicates a high or a low probability that the targeted

Author for correspondence: Stephanie M. van Rooden, Email: S.M.vanRooden@umcutrecht.nl

PREVIOUS PRESENTATION: Results of this study (oral presentation no. O1171) were presented at the 29th European Conference on Clinical Microbiology and Infectious Diseases on April 16, 2019, in Amsterdam, The Netherlands.

Cite this article: van Rooden SM, *et al.* (2020). A framework to develop semiautomated surveillance of surgical site infections: An international multicenter study. *Infection Control & Hospital Epidemiology*, 41: 194–201, <https://doi.org/10.1017/ice.2019.321>

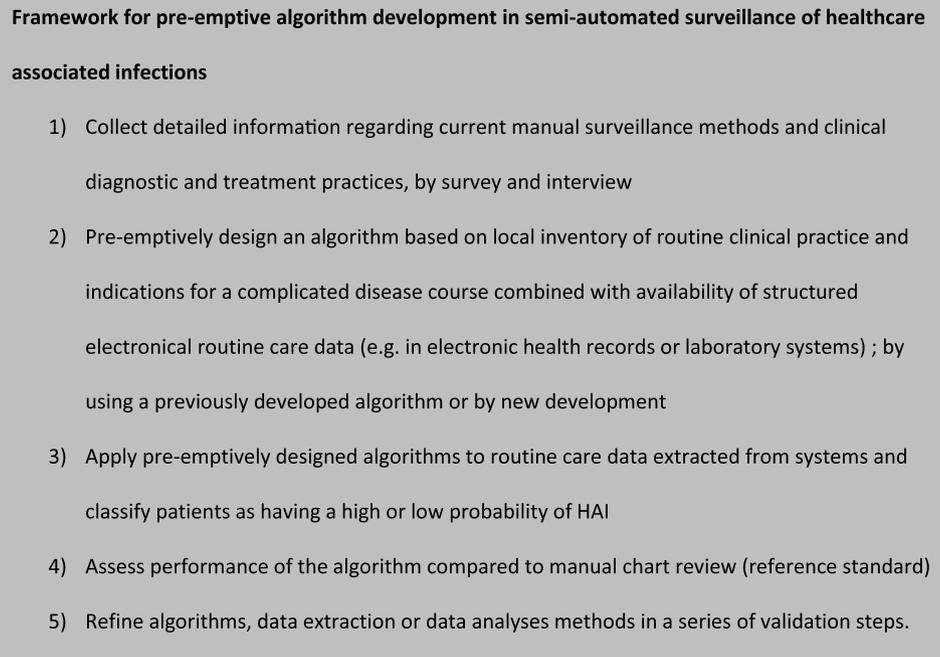


Fig. 1. A framework for pre-emptive algorithm development in semiautomated surveillance of healthcare-associated infections.

HAI occurred. Only procedures classified as high probability undergo manual chart review to confirm the presence of an HAI. This automated preselection increases standardization and reduces the number of charts requiring manual review by >75% while retaining the possibility of interpreting data on clinical signs and symptoms.⁸ In the past decade, multiple publications have described automated surveillance with good performance,^{9–12} but the automation of surveillance has not yet been adopted widely.¹³ Extension to other hospitals may have been hampered by previous efforts being restricted to single hospitals with differences in information technology (IT) systems, data availability, and local diagnostic and treatment practices, or by relying on data-driven algorithm development. Absence of guidance on how to develop, implement, and maintain such systems limits their widespread adoption. A framework to develop algorithms for semiautomated surveillance, applicable to routine care data and not requiring complex data-driven modeling, may facilitate the development of reliable algorithms implementable on a larger scale.

In this observational, international multicenter cohort study, we assessed the performance and feasibility of such a framework for pre-emptive algorithm development for semiautomated surveillance (Fig. 1) applied to the surveillance of deep incisional and organ-space SSIs after hip and knee arthroplasty (THA and TKA), cardiac surgery, and colon surgery. In addition, we evaluated the generalizability of algorithms between hospitals.

Methods

In this retrospective cohort study, we assessed the performance of a framework for the development of semiautomated surveillance by comparing the developed algorithms to routinely performed traditional manual surveillance for the detection of deep incisional and/or organ-space SSI (reference standard), according to locally applied definitions. We focused on SSIs after THA and TKA, cardiac surgery, and colon surgery because these are high-volume procedures commonly targeted by surveillance programs.

Hospital and patient inclusion

European hospitals participating in the CLIN-NET network¹⁴ were recruited for this study. We considered only hospitals with long-term experience with manual SSI surveillance after THA and TKA and cardiac surgery, and (optionally) other surgical procedures, from 2012 onward. In each hospital, the feasibility of data extraction from the EHR for the same period was assessed. Dates of (re)admission and discharge, surgical records, microbiology results, and in-hospital mortality were the minimum requirements for participation. Optional data included demographics, antibiotic prescriptions, outpatient clinic visits, temperature, other vital signs, radiology ordering, clinical chemistry, and administrative data. Data collection was performed retrospectively between July 2017 and May 2019. The targeted number of SSIs during the surveillance period had to be 60–80 per center to estimate performance with sufficient precision. All procedures included in the participating hospitals' targeted surveillance were included nonselectively in the study. From all centers, local approval and waivers of informed consent were obtained.

Automated surveillance framework

The framework was applied to each hospital by the coordinating center (University Medical Center Utrecht, The Netherlands) and consisted of the following steps:

Step 1: Inventory of local surveillance and clinical procedures and availability of electronic routine care data. Centers completed a questionnaire for each targeted procedure collecting information on (1) local surveillance procedures (ie, SSI definitions, selection of the surveillance population), (2) clinical procedures regarding standard of care and diagnostic and therapeutic practice in case of SSI suspicion, and (3) the availability of relevant electronically routine care data (see the supplementary material online: [Questionnaire SSI Surveillance Methods](#)). If needed, hospitals were contacted for a follow-up discussion.

Table 1. Overview of Surveillance Population (Reference Data)

Surgical Procedure in Surveillance	Hospital	Surgery Selection	Applied Definitions	Follow-up Period in Surveillance, ^d	Study Period	No. of Procedures in Surveillance, No. (No. SSIs, % SSIs)	No. of Procedures in Surveillance and in Extracted Routine Care Data (Reference Data), No. (No. SSIs, % SSIs) ^a
Hip/knee arthroplasty	A	Specified selection of procedures ^b	National	90	2015–2016	1,525 (8, 0.5)	1,509 (8, 0.5)
	B	Elective primary TKA/ THA	CDC	365	2016 ^c 2015–2016	334 (6, 1.8) 707 (11, 1.6)	326 (6, 1.8) 686 (11, 1.6)
	C	Specified selection of procedures, ^b no prior procedure <6 months	National	365	2012–2016	2,575 (19, 0.7)	2,575 (19, 0.7)
Cardiac surgery	A	Specified selection of procedures, ^b	National	90	2015–2016	2,645 (34, 1.3)	2,333 (33, 1.4)
	B	CABG and valve replacement	CDC	30/365 ^d	2016 ^c 2015–2016	512 (11, 2.1) 850 (18, 2.1)	440 (9, 2.0) 725 (15, 2.1)
	C	Specified selection of procedures, ^b no prior procedure 6 mo	National	30/365 ^d	2012–2016	1,989 (46, 2.3)	1,989 (46, 2.3)
Colon surgery	A	Specified selection of procedures, ^b no prior procedure <30 d	National	30	2013–2016	1,325 (100, 7.5)	1,293 (89, 6.9)
	B	Elective colon/rectal surgeries	CDC	30	2016 ^c 2015–2016	208 (7.7) 396 (43, 10.6)	200 (16, 8.0) 386 (43, 11.4)
	C	Specified selection of procedures, ^b no prior procedure <6 mo	National	30	2012–2016	2,227 (98, 4.4)	2,227 (98, 4.4)

Note. SSI, deep incisional or organ-space surgical site infection; TKA, total knee arthroplasty; THA, total hip arthroplasty; CDC, Centers for Disease Control and Prevention²⁴; CABG, coronary artery bypass grafting.

^aProcedures included in surveillance that could be matched to routine care data extracted from electronic health records; this served as a reference population for comparison of the algorithm performance.

^bThe selection of surgeries (related to hip and knee prosthesis, cardiac surgery, or colon surgery) consists of a large number of procedure codes, not further specified in this table.

^cSeparate analyses were performed for hospital B for surveillance of 2016 and 2015–2016, because of data availability of data on antibiotics starting 2016.

^dFollow-up period 30 days for procedures without foreign material, for procedures with foreign material 365 days.

Step 2: Algorithm design. Algorithms were pre-emptively designed based on the information collected in the inventory (step 1). Algorithms classified surgical procedures as having a high probability of SSI if they met the criteria of different components representing different possible indicators of HAI, adapted to targeted procedures. With data selection and defining component criteria, the following considerations were taken into account: (1) the relevance of data to serve as an HAI indicator for hospitals; (2) data availability across hospitals; (3) robustness to small changes in clinical practice or documentation; (4) ease of application irrespective of local IT systems and epidemiological support; and (5) a primary focus on optimizing sensitivity (ie, a high detection rate) followed by positive predictive value (PPV) because the framework is set up for semiautomated surveillance. For each targeted procedure, both a standardized algorithm based on common clinical practices across all hospitals and a center-specific algorithm adapted to specific local clinical procedures were developed. If available, previously developed algorithms were used as the standardized algorithm and were validated.

Step 3: Classification high or low probability of SSI. Algorithms were applied to the data extracted from each center to classify procedures as high or low probability of an SSI.

Steps 4 and 5: Assessing and refining algorithm performance. Results of the semiautomated algorithms were compared to traditional surveillance to determine accuracy and efficiency of SSI detection. Subsequently, based on group-level analysis of results per algorithm component (without evaluating individual

procedures), and discussion with each hospital, structurally missing data or miscoding were corrected. Second, a detailed case-by-case discrepancy analysis was performed to obtain insight into reasons for misclassification. All false-negative cases, and a selection of false-positive and concordant cases, were reassessed by each center, blinded for algorithm outcome. If needed, additional corrections in algorithm application were made. Errors in manual surveillance potentially discovered after reassessment were not reclassified.

Analyses

The primary end point of this study was the determination of the accuracy and efficiency of SSI detection by the semiautomated algorithms, as measured by sensitivity, PPV, and workload reduction as proportion of charts requiring manual review, compared to traditional manual surveillance. Analyses were performed at the procedural level. In addition, discrepancy analyses explored reasons for misclassification. Finally, the results of this study provide an indication of the feasibility of broad adoption of this semiautomated surveillance framework.

Results

Inclusion of hospitals and surveillance characteristics

Initially, 4 hospitals were recruited. However, the required EHR data extraction for 1 hospital was not possible, and this hospital was

Table 2. Data Collection

Data Category	Hospital A	Hospital B	Hospital C
Demographics	Available	Available	Available
Hospital mortality	Available	Available	Available
(Re-)admission & discharge	Available	Available	Available
Operating records	Available	Available^a	Available
Microbiology data	Available	Available >2014^a	Available
Antibiotic use in ICU	Available	Not available	Not available
Antibiotic use in general ward	Available	Available >2015	Not available
Outpatient clinic visits	Available	Available, not collected	Not available
Temperature & vital signs	Available	Available, not collected	Not available
Clinical chemistry results	Available	Available, not collected	Available, not collected
Radiology ordering	Available >2012	Available	Available, not collected

Note. For bolded rows, variables are considered minimally required; only hospitals with these variables available were included in the study for testing the framework.

^aAppeared not fully complete during analyses, but could no longer be extracted due to change of systems.

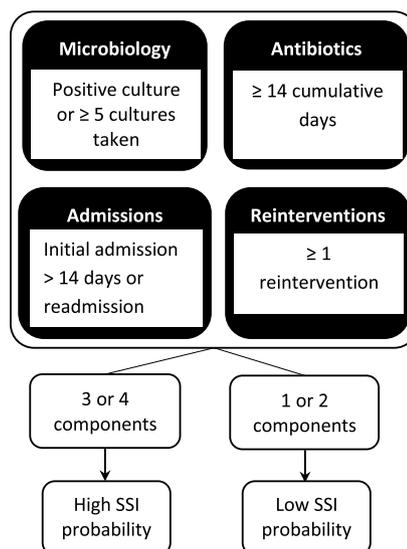
therefore excluded from further analyses. The included hospitals were Amphia hospital (Breda, Netherlands), Dupuytren University Hospital (Limoges, France), the Bellvitge University Hospital (Barcelona, Spain), referred to as hospital A-C (random order).

Table 1 provides an overview of the surveillance population, definitions used in manual surveillance, and the number of procedures that could be linked to EHR data and, hence, to serve as reference data to evaluate algorithm performance. In total, 4,770 THA and TKA procedures, 5,047 cardiac surgeries, and 3,906 colon surgeries, with 38 SSIs for THA and TKA, 94 SSIs for cardiac surgeries, and 230 SSIs for colon surgeries. An overview of available EHR data is presented in Table 2. Because data for antibiotics from hospital B were only available for 2016, the framework was applied to data without antibiotics and additionally to a subset of data including antibiotics.

Algorithm development

All pre-emptively developed classification algorithms included components reflecting microbiological culture results, admissions (prolonged length of stay or readmission), and reinterventions, with criteria adapted to the algorithm (standardized or center specific) and targeted procedure. Additionally, algorithms including antibiotic prescriptions or radiology ordering components were defined to accommodate the differences between centers in data availability. Procedures were classified as high probability of SSI if a patient scored positive on a combination of components, including a mandatory component for some algorithms. For THA and TKA, the algorithm developed by Sips et al¹⁵ was applied in hospitals when data on antibiotics were available. A schematic overview of the algorithms for THA and TKA is provided in Figure 2 as an example. Detailed algorithm descriptions and a flowchart of framework application are provided in the supplementary material online.

(a) Algorithm when data on antibiotics is included



(b) Algorithm when data on antibiotics is not included

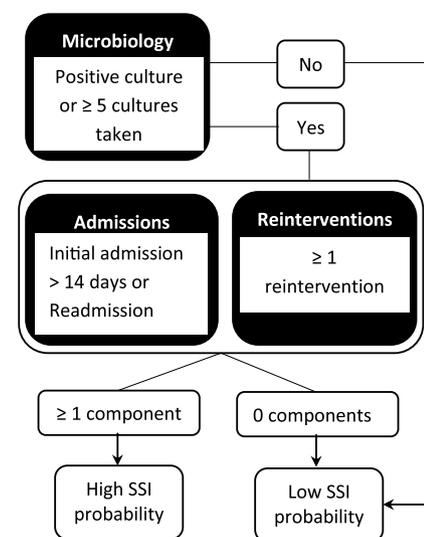


Fig. 2. Schematic overview of developed algorithm (example hip and knee replacement, standardized algorithm). Example schematic overview of algorithms developed for semi-automated surveillance after hip and knee arthroplasty, with components standardized for all centers (standardized algorithm). Only high SSI probability cases underwent manual chart review. Cultures were restricted to relevant body sites and excluded cultures taken prior to day 1. Readmissions were restricted to admission by the operating specialty. All reinterventions performed by initial specialty were included. Antibiotics included all prescriptions with ATC code J01. In the algorithms designed specifically for each center, the microbiology component was defined as a positive culture; length of stay, reinterventions, and antibiotics were specified according to each hospital's clinical procedures. Details of standardized and center-specific algorithms for all interventions are reported in the supplementary material online.

Algorithm performance

Table 3 presents the performance of all algorithms prior to case-by-case discrepancy analyses, and the results are presented in Table 4. Overall, the performance of the standardized algorithms was high in terms of sensitivity and workload reduction. Center-specific algorithms often achieved higher workload reduction, but at the cost of sensitivity.

Table 3. Performance of Algorithms for Semiautomated Surveillance of Deep Incisional and Organ-Space Surgical Site Infections

Surgical Procedure	Antibiotics Included Algorithm	Hospital	Standardized Algorithm, % (No./Total) ^a			Center-Specific Algorithm, % (No./Total) ^a		
			Sensitivity ^b	PPV ^c	Workload Reduction ^d	Sensitivity ^b	PPV ^c	Workload Reduction ^d
Hip/knee prosthesis	Antibiotics	A	100.0 (8/8)	17.4 (8/47)	96.9 (47/1,509)	100.0 (8/8)	20.0 (8/40)	97.3 (40/1,509)
		B	83.3 ^e (5/6)	62.5 (5/8)	97.5 (8/326)	50.0 (3/6)	37.5 (3/8)	97.5 (8/326)
	No antibiotics data	B	81.8 ^e (9/11)	42.9 (9/21)	96.9 (21/686)	81.8 ^e (9/11)	9.8 (9/92)	86.6 (92/686)
		C	94.7 ^e (18/19)	18.4 (18/98)	96.2 (98/2,575)	94.7 ^e (18/19)	15.1 (18/119)	96.2 (119/2,575)
Cardiac surgery	Antibiotics	A	97.0 (32/33)	34.8 (32/92)	96.1 (92/2,333)	93.9 (31/33)	43.7 (31/71)	97.0 (71/2,333)
		B	66.7 (6/9)	19.4 (6/31)	93.0 (31/440)	44.4 (4/9)	33.3 (4/12)	97.3 (12/440)
	No antibiotics data	B	100.0 (15/15)	7.9 (15/191)	73.7 (191/725)	93.3 (14/15)	19.7 (14/71)	90.2 (71/725)
		C	95.7 ^e (44/46)	8.3 (44/531)	73.2 (531/1,989)	89.1 (41/46)	21.5 (41/191)	90.4 (191/1,989)
Colon surgery	Antibiotics and radiology ordering included	A	93.3 (83/89)	36.1 (83/230)	82.2 (230/1,293)	86.5 (77/89)	45.3 (77/170)	86.9 (170/1,293)
		B	100.0 (16/16)	30.2 (16/53)	73.6 (53/201)	56.3 (9/16)	42.9 (9/21)	89.6 (21/201)
	Antibiotics and radiology ordering not included	B	83.7 (36/43)	33.6 (36/107)	72.3 (107/386)	48.8 (21/43)	43.8 (21/48)	87.6 (48/386)
		C	93.9 (92/98)	16.6 (92/554)	75.1 (554/2,227)	76.5 (75/98)	27.9 (75/269)	87.9 (269/2,227)

Note. PPV, positive predictive value.

^aA detailed description of applied algorithms is provided in the supplementary material online.

^bSensitivity is defined as the number of procedures with SSI in the reference surveillance data that were classified by the algorithm as high probability of an SSI.

^cThe positive predictive value is defined as the number of SSI in the reference surveillance data within all procedures that were classified as high probability by the algorithm.

^dThe workload reduction is defined as the number of procedures that require manual confirmation (ie procedures that were indicated by the algorithm with a high probability of an SSI) as compared to all procedures in traditional surveillance.

^eResults discrepancy analyses: reclassification of false negative cases: 100% case detection if corrected.

Standardized algorithms for surveillance of SSIs after THA and TKA had a sensitivity ranging from 81% to 100% across hospitals. Upon reconsideration, all missed SSIs were deemed “no SSI” in the discrepancy analyses; thus, 100% of cases could be detected. A workload reduction of >95% was achieved in all centers. The center-specific algorithms yielded a sensitivity ranging from 50% to 100% and a workload reduction ranging from 87% to 98% across hospitals.

For cardiac surgery, the sensitivity of the standardized algorithms appeared nearly perfect, with a 73%–96% workload reduction, except for the algorithm including antibiotics in hospital B. Of 9 SSIs, 3 were missed; 1 SSI was reconsidered by the center as “no SSI” in the discrepancy analyses. The sensitivity of the center-specific algorithms across centers ranged from 44% to 95%, with a 90%–97% workload reduction.

The standardized algorithm for SSI detection after colon surgery showed >90% sensitivity in all centers, but it was lower for the algorithm without antibiotics in hospital B. All centers achieved a workload reduction between 72% and 82%. Sensitivity of the center-specific algorithms ranged from 49% to 82% across centers.

No formal comparisons were performed, but standardized algorithms with and without an antibiotics component showed comparable overall sensitivity. The PPV and workload reduction were better in algorithms including an antibiotic component for cardiac surgery only. To further assess generalizability of this finding, the performance of standardized algorithms without an antibiotics component was evaluated for hospital A as a sensitivity analysis (see the supplementary material on line: Detailed Algorithm Description, flowchart A). This analysis yielded the following results: for THA and TKA, sensitivity of 100%, a PPV 19%, and a workload reduction of 97%; for cardiac procedures, sensitivity of 94%, a PPV of 18%, and a workload reduction of 93%; for colon surgery, sensitivity of 93%, PPV of 33%, and a workload reduction of 80%.

For all targeted procedures, most SSIs missed by the standardized algorithms (ie, false negatives) were reassessed as “no SSI” in the discrepancy analyses; hence, they were correctly classified by the algorithm. For colon surgery, reasons for missed SSIs included not fulfilling the mandatory microbiology component (algorithms without antibiotics) and incomplete extraction of microbiology data (hospital B) and procedures (ie, drainage or debridement; all hospitals). The SSIs missed by the center-specific algorithms could be explained by the component criteria being too specific (eg, antibiotics after THA/TKA and cardiac surgery in hospital B and microbiology after colon surgery in hospital C). The main reasons for falsely identified high-probability cases were errors in the reference data (SSI after reconsideration in the discrepancy analyses), superficial SSIs or other complications, and patients with pre-existing infection being included in the surveillance population.

Feasibility of framework application

Application of this semiautomated surveillance framework was feasible in the 3 hospitals included in the study; algorithms with good performance were developed, applied, and validated. The questionnaire (framework step 1) provided sufficient information for pre-emptive algorithm development, although application of the algorithm required collecting further details regarding selection of data and technical specification from IT specialists, infection control practitioners, and clinicians. The fourth hospital had to be excluded from this study because historical data extraction was impossible for most data despite considerable effort. In hospital B, data on microbiology results and reinterventions could not be completed due to changes in the IT system. Factors enhancing the feasibility of this framework as encountered during this study are presented in Table 5.

Table 4. Discrepancy Analyses: Reasons for Discrepancies Between the Algorithm Results and Results of Manual Surveillance (Reference), for false-negative (SSIs Missed by Algorithm) and False-Positive Cases

Algorithm and Hospital	Hip and Knee Replacement								Cardiac Surgery								Colon Surgery							
	Standardized Algorithm				Specific Algorithm				Standardized Algorithm				Specific Algorithm				Standardized Algorithm				Specific Algorithm			
	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)	A (Incl AB)	B (Incl AB)	B (No AB)	C (No AB)
Total false-negative cases, no.	...	1	2	1	...	3	2	1	1	3	...	2	2	5	1	5	6	...	7	6	12	7	22	23
Reasons for false negatives: Missed SSIs (count)																								
Reclassification of reference data: true negative (“no SSI”), no.		1	2	1		1	2	1		1		2		2	1	4			1	3			1	8
Field or code selection incomplete, no.									1								5		2	2	5	3	11	
Incomplete data extraction, no.										2			1	2								7	10	
SSI diagnosis not based on microbiology, no.																1	6		4	2	12		7	5
Algorithm not matching clinical procedures, no.										2									1	1				
Too specific definition algorithm component, no.						2 ^{a,b}								2 ^a										10 ^c
Relevant information stored as free text, no.																	1	1			1			15
Diagnostics in other hospital, no.													1											
Total false-positive cases, no.	17	1	4	3	17	1	4	4	7	1	11	6	7	...	11	6	9	5	10	8	9	5	10	8
Reasons for false positives (count)																								
Reclassification of reference data: true positive, no.	11	1	2	1	11		1	1			1	2			1	2	2	1	3	4	2	1	3	4
Superficial SSI, no.									2		3		2		3			1	3	3		1	3	3
Pre-existent SSI, no.	2			1	2			1	1			1	1			1								
Other complications, no.	4		2	1	4	1	3	1	4	1	4	3	4		4	3	7	3	4		7	3	4	1
No signs of infection or antibiotics, no.								1				3			3									

Note. SSI, surgical site infection, AB, antibiotics component of algorithm. Ellipses (...) indicate no false negatives, all cases detected and no false positives included. Multiple reasons may be applicable to 1 case.

^aAntibiotic component.

^bReintervention component.

^cMicrobiology component.

Table 5. Factors Enhancing Successful Framework Application

Enhancing Factor	Motivation
Early involvement and collaboration between clinicians, infection control practitioners and IT specialists	Knowledge of different fields of expertise need to be combined: Early involvement of IT specialists was important to learn what (historical) data is feasible to extract in a structured form, from various systems. Clinicians and infection control practitioners have knowledge on relevance of data and of registration practices, such as fields or codes that are applied.
Access to data of historical systems	Validation required the use of historical data; changing electronic health records systems can result in loss of access to essential data.
Data management support	In all hospitals multiple systems for registration of routine care data were running in parallel. IT, medical intelligence specialists, and data managers facilitated algorithm application.
Validation of extracted data	For correct application of the algorithm, validation of data quality is important (corrections had to be made to correct structurally missing data).
High quality reference data for algorithm validation	In the discrepancy analyses, cases in the manual surveillance were reclassified (SSI to “no SSI” and vice versa), limiting performance estimation of the algorithms

Note. Enhancing factors increase the feasibility of framework application, based on lessons learned during the framework application.

Discussion

This retrospective study assessed a framework for the development of semiautomated HAI surveillance in 3 different European hospitals, focusing on deep incisional and organ-space SSIs after THA and TKA, cardiac procedures, and colon surgery. This framework achieved a high detection rate of SSI (sensitivity) as well as a 72%–98% reduction of manual chart review workload when it was applied in hospitals that differed with respect to surveillance population, HAI definition, and clinical procedures. Because this method of pre-emptive algorithm development relies on a limited number of data sources that can often be extracted from the EHR and does not require complex modeling, calibration, or dealing with missing data,¹⁶ we expected it to be accessible to many hospitals.

The algorithms, developed without any prior knowledge other than information obtained from the survey and interviews, performed well. The ‘standardized algorithms’ developed with the purpose of being applicable in multiple hospitals had high sensitivity while achieving a substantial workload reduction, given that data extraction was complete. These algorithms offer possibilities for larger-scale standardization of semiautomated surveillance. Although algorithms with a more specific component definition, tailored to each center’s characteristics, often achieved a higher positive predictive value, the gain in workload reduction was limited and did not offset the loss in sensitivity. Components appeared to be too specifically defined, either because clinical procedures in case of an SSI were less standardized than anticipated or because the reasoning for SSI diagnosis

was less straightforward. Furthermore, the standardized algorithm components are likely more robust to changes in clinical practice over time and are easier to implement and maintain than specific definitions. Hence, for the purpose of semiautomated surveillance, algorithms with more generally defined components are preferable.

The algorithm previously developed for semiautomated surveillance of SSI after TKA and THA¹⁵ was validated in 2 hospitals in this study; all SSIs were detected and a similar reductions in workload were achieved, although the number of SSIs was limited. The algorithm without antibiotics performed similarly. This latter observation also held true for the other types of surgery. In previous studies, data on antibiotics were added to enhance case findings,^{17–19} but this addition was not essential to this study, and the same held true for data on radiological interventions. Although additional information could be used to optimize algorithm performance, it is feasible to develop well-performing algorithms that rely solely on microbiology results, admission and discharges dates, and procedures codes. These algorithms may be broadly adoptable because extraction of this information in an easy-to-process format was possible in all hospitals except one (due to impossibilities of accessing historical data in legacy EHR), and no computation is required (eg, deriving changes in clinical chemistry results). These elements do not depend on interpretation by medical coders; therefore, they are potentially more reliable than, for example, billing codes.^{12,13}

In algorithms without data on antibiotic prescriptions, fulfilling the microbiology component was mandatory to limit the number of false-positive cases. Based on the survey (framework step 1), we anticipated that cultures were taken whenever deep incisional or organ-space SSI was suspected; however, the discrepancy analyses revealed that the SSI determination was not always based on microbiology. The appropriateness of a mandatory component can be discussed because culture-negative infections can also meet the definition of deep incisional or organ-space SSI.²⁰ More flexibility could have increased the sensitivity, but there is always a trade-off between sensitivity and PPV. Accepting missed cases could be another alternative as long as comparability is ensured.²¹

Application of this semiautomated surveillance framework was feasible, but this study revealed several important conditions for success. Early involvement of IT specialists was essential because they provided an overview of available data and because the complexity of data management depends on IT support or the medical intelligence department. Equally important is knowledge of local clinical procedures and registration practices to understand what fields or codes should be used in analyses. Hence, data extraction requires close collaboration between infection control practitioners, clinicians, and IT specialists.

The discrepancy analyses revealed procedures misclassified by manual surveillance, which highlights the importance of a validated reference surveillance population when testing the performance of models. Because the selection of cases for the discrepancy analyses was nonrandom, the results of reclassification could not be reliably extrapolated to the entire population and performance estimates were not recalculated. Hence, the performance of the algorithm was likely underestimated.

Our results have some limitations. Even though this study included hospitals from different countries that varied in applied HAI definitions, procedures in surveillance, and clinical and surveillance practices, the generalizability of the results of this framework to other centers may be limited. Only 3 hospitals were

included in this study, all from Western Europe. Furthermore, differences in the availability of high-quality data for algorithm application (eg, microbiology results) or clinical registration practices will likely have an impact on the applicability of the presented algorithms in new settings, although similar steps in algorithm development could be undertaken. Performance estimation was further limited by a lower total number of SSIs than anticipated and by incomplete data extraction; both are consequences of limited availability of EHR data. Also, we did not measure net time in workload reduction; this cannot be estimated as a linear reduction proportionate to the charts needed for review. Because implementation of semiautomated surveillance was outside the scope of this study, no guidance on investments in human resources and material were developed in this study. More detailed practical guidance could be obtained from implementation studies.

This study was a proof-of-principle investigation of a framework for semiautomated HAI surveillance algorithm development that has promise for broader implementation. Algorithms with good performance can be developed without the need for specific modeling by each hospital and based on limited data sources only. Further validation could provide insight into the feasibility of broader applications of this method, both in other hospitals and for other targeted HAIs. With standardized, semiautomated surveillance on a larger scale, the number of surveyed procedures can be expanded to facilitate local quality improvement, (inter)national comparisons, or outcome measurements in clinical trials.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/ice.2019.321>

Acknowledgments. We thank Erik Ribbers, Miranda van Rijen (Amphia hospital, Breda, Netherlands), Xavier Rodriguez (Bellvitge University Hospital, Barcelona, Spain), and the infection prevention departments, IT departments, and surgical departments of all participating hospitals for their valuable contribution to this project.

Financial support. This research was supported by the EPI-Net COMBACTE-MAGNET project, funded by the Innovative Medicines Initiative Joint Undertaking (grant no. 115737), the resources of which included a financial contribution from the European Union Seventh Framework Programme (grant no. FP7/2007–2013) and the European Federation of Pharmaceutical Industries and Associations companies' in-kind contribution.

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

References

- European Centre for Disease Prevention and Control. Healthcare-associated infections: surgical site infections. In: *Annual Epidemiological Report for 2016*. Stockholm: ECDC; 2018.
- Cassini A, Plachouras D, Eckmanns T, *et al*. Burden of six healthcare-associated infections on European population health: estimating incidence-based disability-adjusted life years through a population prevalence-based modelling study. *PLoS Med* 2016;13:1–16.
- Geubbels ELPE, Nagelkerke NJD, Mintjes-De Groot AJ, Vandenbroucke-Grauls CMJE, Grobbee DE, De Boer AS. Reduced risk of surgical site infections through surveillance in a network. *Int J Qual Heal Care* 2006;18:127–133.
- Brandt C, Sohr D, Behnke M, Daschner F, Rüden H, Gastmeier P. Reduction of surgical site infection rates associated with active surveillance. *Infect Control Hosp Epidemiol* 2006;27:1347–1351.
- Abbas M, de Kraker MEA, Aghayev E, *et al*. Impact of participation in a surgical site infection surveillance network: results from a large international cohort study. *J Hosp Infect* 2019;102:267–276.
- Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA* 2011;305:2462–2463.
- Schröder C, Behnke M, Gastmeier P, Schwab F, Geffers C. Case vignettes to evaluate the accuracy of identifying healthcare-associated infections by surveillance persons. *J Hosp Infect* 2015;90:322–326.
- Van Mourik MSM, Perencevich EN, Gastmeier P, Bonten MJM. Designing surveillance of healthcare-associated infections in the era of automation and reporting mandates. *Clin Infect Dis* 2018;66:970–976.
- Russo PL, Shaban RZ, Macbeth D, Carter A, Mitchell BG. Impact of electronic healthcare-associated infection surveillance software on infection prevention resources: a systematic review of the literature. *J Hosp Infect* 2018;99:1–7.
- Freeman R, Moore LSP, García Álvarez 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.
- Sips ME, Bonten MJM, van Mourik MSM. Automated surveillance of healthcare-associated infections. *Curr Opin Infect Dis* 2017;30:425–431.
- Gastmeier P, Behnke M. Electronic surveillance and using administrative data to identify healthcare associated infections. *Curr Opin Infect Dis* 2016;29:394–399.
- 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. *J Am Med Inf Assoc* 2014;21:942–951.
- CLIN-Net website. <https://www.combacte.com/about/clin-net/>. Accessed June 18, 2019.
- Sips ME, Bonten MJM. Semiautomated surveillance of deep surgical site infections after primary total hip or knee arthroplasty. 2017;38:732–735.
- Van Mourik MSM, Troelstra A, Van Solinge WW, Moons KGM, Bonten MJM. Automated surveillance for healthcare-associated infections: opportunities for improvement. *Clin Infect Dis* 2013;57:85–93.
- Yokoe DS, Khan Y, Olsen MA, *et al*. Enhanced surgical site infection surveillance following hysterectomy, vascular, and colorectal surgery. *Infect Control Hosp Epidemiol*. 2012;33:768–773.
- Bolon MK, Hooper D, Stevenson KB, *et al*. Improved surveillance for surgical site infections after orthopedic implantation procedures: extending applications for automated data. *Clin Infect Dis* 2009;48:1223–1229.
- Perdiz LB, Yokoe DS, Furtado GH, Medeiros EAS. Impact of an automated surveillance to detect surgical-site infections in patients undergoing total hip and knee arthroplasty in Brazil. 2017;37:2016–2018.
- Chalfine A, Cautet D, Lin WC, *et al*. Highly sensitive and efficient computer-assisted system. *Infect Control Hosp Epidemiol* 2006;27:794–801.
- Hota B, Lin M, Doherty JA, *et al*. Formulation of a model for automating infection surveillance: algorithmic detection of central-line-associated bloodstream infection. *J Am Med Informatics Assoc* 2010;17:42–48.
- Karami M, Rahimi A, Shahmirzadi AH. Clinical data warehouse. *Health Care Manag (Frederick)* 2017;36:380–384.
- Woeltje KF, Lin MY, Klompas M, Oliver Wright M, Zuccotti G, Trick WE. Data requirements for electronic surveillance of healthcare-associated infections. *Infect Control Hosp Epidemiol* 2017;35:1083–1091.
- Surgical site infection event. Centers for Disease Control and Prevention website. <https://www.cdc.gov/nhsn/pdfs/pscmanual/9pscscscurrent.pdf>. Published 2019. Accessed November 4, 2019.