



Original Article

Hospital-acquired infections surveillance: The machine-learning algorithm mirrors National Healthcare Safety Network definitions

Stephani Amanda Lukasewicz Ferreira RN, MSc¹ , Arateus Crysham Franco Meneses BSc¹,
Tiago Andres Vaz PhD¹ , Otavio Luiz da Fontoura Carvalho PharmaD¹, Camila Hubner Dalmora MD¹,
Daiane Pressotto Vanni RN², Isabelle Ribeiro Berti MD² and Rodrigo Pires dos Santos MD, PhD¹

¹Qualis, Porto Alegre, Rio Grande do Sul, Brazil and ²Tacchini Hospital, Bento Gonçalves, Rio Grande do Sul, Brazil

Abstract

Background: Surveillance of hospital-acquired infections (HAIs) is the foundation of infection control. Machine learning (ML) has been demonstrated to be a valuable tool for HAI surveillance. We compared manual surveillance with a supervised, semiautomated, ML method, and we explored the types of infection and features of importance depicted by the model.

Methods: From July 2021 to December 2021, a semiautomated surveillance method based on the ML random forest algorithm, was implemented in a Brazilian hospital. Inpatient records were independently manually searched by the local team, and a panel of independent experts reviewed the ML semiautomated results for confirmation of HAI.

Results: Among 6,296 patients, manual surveillance classified 183 HAI cases (2.9%), and a semiautomated method found 299 HAI cases (4.7%). The semiautomated method added 77 respiratory infections, which comprised 93.9% of the additional HAIs. The ML model considered 447 features for HAI classification. Among them, 148 features (33.1%) were related to infection signs and symptoms; 101 (22.6%) were related to patient severity status, 51 features (11.4%) were related to bacterial laboratory results; 40 features (8.9%) were related to invasive procedures; 34 (7.6%) were related to antibiotic use; and 31 features (6.9%) were related to patient comorbidities. Among these 447 features, 229 (51.2%) were similar to those proposed by NHSN as criteria for HAI classification.

Conclusion: The ML algorithm, which included most NHSN criteria and >200 features, augmented the human capacity for HAI classification. Well-documented algorithm performances may facilitate the incorporation of AI tools in clinical or epidemiological practice and overcome the drawbacks of traditional HAI surveillance.

(Received 21 March 2023; accepted 16 September 2023)

Hospital-acquired infection (HAI) surveillance is the cornerstone of infection prevention. In the United States, HAIs are responsible for 72,000 preventable deaths each year.¹ During the coronavirus disease 2019 (COVID-19) pandemic, we have witnessed increases in most HAI rates.² In Brazil, the criteria for HAI surveillance is defined by the National Health Surveillance Agency (Anvisa, the Agência Nacional de Vigilância Sanitária), which are mainly based on the National Healthcare Safety Network (NHSN) criteria for infection classification.^{3,4}

Most often, the search for HAI is conducted manually by well-trained infection control professionals (ICPs). However, this process may suffer from interrater variability, depending on ICP expertise, and is time-consuming. Furthermore, the chosen method, global or device/culture-focused surveillance, for example, may produce variable results.⁵

Recently, automated and semiautomated methods that use data from electronic health records (EHR) have been developed. In semiautomated methods, the possible infections above a threshold are shown and an individual evaluates these infections and confirms or discards the diagnosis.⁶ These methods have shown good performance for HAI classification. However, they show variable performance in sensitivity, specificity, predictive values, and accuracy.^{7,8} Factors like the hospital setting, type of infection surveilled, and patient population analyzed may contribute to these variations.⁸

In a previous study, in which we searched for infections globally, the machine learning (ML) model outperformed manual surveillance by 42%, and time spent on record review decreased by 71%.⁹ In this study, we compared manual surveillance with the supervised ML semiautomated method, and we explored the types of infection that each type of surveillance identified as well as the features of importance depicted by the ML model.

Methods

Tacchini Hospital is a general hospital with 251 beds that provides clinical and surgical care in southern Brazil. It serves ~400,000 people and includes clinical specialties and surgical procedures for

Corresponding author: Stephani Amanda Lukasewicz Ferreira, RN, MSc, Qualis, 1022 Osvaldo Aranha Ave, Room 1101, Porto Alegre, RS, 90035-191, Brazil. E-mail: stephani@portalqualis.com.br

Cite this article: Lukasewicz Ferreira SA, Franco Meneses AC, Vaz TA, *et al.* Hospital-acquired infections surveillance: The machine-learning algorithm mirrors National Healthcare Safety Network definitions. *Infect Control Hosp Epidemiol* 2024. doi: [10.1017/ice.2023.224](https://doi.org/10.1017/ice.2023.224)

general and pediatric surgery, gynecology, mastology, obstetric surgeries, oncology, neurology, orthopedics, plastic, vascular, and urology surgeries. During the study period, the hospital had 2 adult ICUs totaling 30 beds: 26 beds for clinical and four for surgical patients. The infection control team, composed of a nurse and an infectious disease physician, is responsible for HAI surveillance. Manual surveillance was performed by the hospital infection control team based on bacterial culture results, and patient records were collected for review. Between July and December 2021, a semiautomated surveillance method based on ML algorithms as described by Ferreira *et al*⁹ was implemented in the hospital. In this semiautomated process, the artificial intelligence (AI) tool classifies patients with potential HAIs, and ICPs subsequently validate the classification.

The performance of the ML algorithm has been reported elsewhere.^{9,10} For this study, the manual diagnosis was reviewed by independent specialists and the golden standard was reclassified strictly according to CDC Anvisa criteria. The ML model was retrained according to these new standards. All adult inpatients (aged ≥ 18 years) were included, but individuals treated in outpatient areas were not eligible.

Data queries analyzed by ML models included laboratory results, radiology results (unstructured data), healthcare professional records (unstructured data), antimicrobial prescriptions, vital signs, and invasive procedures. The ML models were based on the random forest algorithm for supervised training because of accuracy and relative ease of explanation. The random forest algorithm randomly uses decision trees to explore potential connections of each variable to an outcome.⁹

The features in the model refer to the individual measurable properties or characteristics of the data that are used as inputs to an ML algorithm. The importance of a feature is the average information gained during the forest construction; it is expressed as a percentage.

The infection control team manually classified the features used in the ML method as follows: concerning signs and symptoms of an infection, patient severity status, variables related to patient comorbidities, antibiotic use, invasive procedures, risk factors for infection, and NHSN- and Anvisa-related criteria. The following features were reclassified or flagged as proxy features: (1) features related to signs and symptoms of infection, antibiotic use, and laboratory culture results and (2) features related to risk factors for infection, patient severity status, invasive procedures, length of stay, patient comorbidities. Considering the data for the entire year of 2021, we ran the ML HAI classification using all features in the algorithm model. We then compared to the performance of risk variables (excluding from the model proxy variables) and proxy variables (excluding from the model the risk variables) using the random forest model in terms of sensitivity, precision (positive predictive value), and accuracy.

This study generated novel data concerning cases identified by the ML algorithm, ranked by probability of HAI. These results highlighted the most important variables or features for the model and facilitated a comparative analysis between manual surveillance and semiautomated surveillance, stratified by HAI classification.

Nonparametric isotonic regression calibration was used to analyze false-positive and false-negative results. For patient comparisons, the χ^2 test was used for categorical variables. Coefficient κ of agreement for categorical outcomes was also used. A 2-tail significance level of 5% was considered, and data analyses were performed using the Statistical Package for the

Table 1. Rate of HAI Classification Comparing the Reference Manual Method and the Additional Infections Identified by Semiautomated ML Algorithm

Variable	Manual Surveillance (N=183), % (n/N)	Additional Infections Identified by Semiautomated ML (N=116), % (n/N)	P Value
UTI	100.0 (52/52)	100.0 (17/17)	...
VAP	95.3 (41/43)	55.5 (11/20)	<.001
Tracheobronchitis	100.0 (16/16)	20.0 (6/30)	...
Non-VAP	100.0 (22/22)	20.0 (5/25)	...
SSI	95.0 (19/20)	43.7 (7/16)	.001
BSI	100.0 (25/25)	100.0 (5/5)	...

Note. HAI, hospital-acquired infection; ML, machine learning; UTI, urinary tract infection; VAP, ventilator-associated pneumonia; SSI, surgical-site infection; BSI, bloodstream infection.

Social Sciences (SPSS) version 16.0 software (IBM, Armonk, NY). This research was registered in Plataforma Brasil (CAAE no. 89737218.0.0000.5305).

Results

During the study period, from 6,296 patients, manual surveillance classified 183 cases (2.9%) of HAI. The semiautomated surveillance found 299 HAI cases (4.7%), 116 more cases (38.7%) than manual surveillance. Both methods identified 180 cases, and the ML tool missed 3 cases detected by the manual review: 1 case of ventilator-associated pneumonia (VAP) and 2 urinary tract infections (UTIs). Considering manual surveillance as the gold standard for infections to be classified as HAIs, κ agreement was 0.991 (95% confidence interval [CI], 0.982–1.0). When we included the infections not classified by manual surveillance, the κ agreement was 0.737 (95% CI, 0.693–0.782). The infection rate (infections per 1,000 patient days) during the manual review was 5.3, and by automated surveillance, the infection rate was 7.5 during the 6-month study period.

Among the 299 cases classified by the ML tool, 68 (22.7%) were urinary tract infections; 63 (21.1%) were ventilator-associated pneumonia (VAP); 48 (16.0%) were tracheobronchitis; 47 (15.7%) were non-VAP; 35 (11.7%) were surgical-site infections (SSIs); 29 (9.6%) were bloodstream infections (BSIs); 4 (1.3%) were abdominal abscess; 2 (0.6%) were peritonitis; 1 (0.3%) was a *Clostridioides difficile* infection; 1 (0.3%) was a skin and soft-tissue infection; and 1 (0.3%) was empyema.

Compared to manual surveillance, the semiautomated method added 77 cases (93.9% of the additional HAIs) in the respiratory infection classification (ie, VAP, non-VAP, and tracheobronchitis). For non-VAP and tracheobronchitis, the numbers of classified cases more than doubled (113.6% and 182.4%, respectively), followed by SSIs (N=15, 66.6%), urinary tract infections (N=17, 32.1%) and bloodstream infections (N=5, 20.8%) (Table 1 and Figure 1).

Regarding the probability of HAIs identified by the algorithm, among those ranked with probability $\geq 90\%$, 79.2% were confirmed. Among HAIs ranked between 80% and 89% probability, 69.2% were confirmed, and among those ranked between 70% to 79%, 25% were confirmed (Table 2). Furthermore, 225 (75.2%) of 299 confirmed infections had a $>70\%$ probability of HAI indicated by the ML algorithm. Comparing manual with semiautomated surveillance, 177 manual surveillance cases (98.3%) had a positive culture result versus 55 (46.2%) from semiautomated surveillance ($P = .001$).

Table 2. The Number of Confirmed Cases Identified by the Machine Learning Algorithm Ranked by Probability of HAI

% Probability of HAI Obtained by ML Algorithm	No. (Confirmed/Suspected Cases)	% of Confirmed Cases
≥90	99/125	79.2
80–89	72/104	69.2
70–79	51/203	25.1
60–69	29/245	11.8
50–59	28/338	8.3
40–59 ^a	15/538	2.8
30–39 ^a	4/3516	0.11

Note. HAI, hospital-acquired infection; ML, machine learning.

^aInfections with probability <50% were reviewed because they were classified as infections by the manual surveillance method.

The ML model considered 447 features or variables for HAI case classification. Among them, 148 features (33.1%) were related to infection signs and symptoms such as fever, secretion, edema, cough, leukocytosis, and pulmonary infiltrate. Furthermore, 101 features (22.6%) were related to patient severity status such as creatinine levels, albumin levels, nutrition status, hemoglobin levels, blood oxygen, blood pressure, and oliguria. Also, 51 features (11.4%) were related to bacterial laboratory results, and 40 features (8.9%) were related to invasive procedures (8.9%; N = 40 features). Finally, 34 features (7.6%) were related to antibiotic use, and 31 features (6.9%) were related to patient comorbidities such as cancer, COVID-19, chronic pulmonary disease, diabetes, stroke, and obesity.

The model characterized features in terms of their importance to HAI classification. The following features were ranked by importance: length of stay (6.0% of relevance), meropenem use (3.7%), intravenous antibiotic use (3.6%), number of medical notes in EHR (2.4%), bacterial culture results (NHSN criteria; 2.2%), number of bacterial culture results (1.9%), patient age (1.7%), patient ventilatory status (1.7%), presence of secretion (NHSN criteria; 1.7%), patient ward transfer (1.4%), the diagnosis of COVID-19 (1.3%), patient being in emergency (1.2%) or intensive care unit (1.1%) ward, number of radiology exams (1.0%), and piperacillin-tazobactam use (1.0%). The following features complete the top 20 most important features identified by the ML model: infection (0.9%), edema (NHSN criteria; 0.9%), fever (NHSN criteria; 0.9%), mechanical ventilation (0.8%) and the number of vital signs registered in the EHR (0.8%).

Considering these 447 features, 229 (51.2%) were similar to those proposed by the NHSN and Anvisa as criteria for HAI classification. When we considered only features related to signs and symptoms of infection, 118 features (79.7%) overlapped with the NHSN and Anvisa definitions of HAI. We also classified 233 features (52.1%) as proxy variables of HAI and 173 features (38.7%) related to infection risk. The performance of the model, comparing all 447 features, proxy, and risk variables, is depicted in Figure 2. Considering the 3 parameters (ie, sensitivity, precision, and accuracy), risk variables performed better than proxy and the overall model had the best performance.

Discussion

Results from ML models are sometimes challenging to interpret and explore, making their implementation in real-life healthcare

situations difficult. In this study, we explored an ML model in its features for HAI classification using the random forest (RF) algorithm. Random forest is called an “off-the-shelf” algorithm; it uses a lot of decision trees for classification (ensemble learning), and it produces understandable prediction rules. A random forest model can handle categorical, continuous, parametric, and nonparametric data, and it is one of the best algorithms for classification of 2 groups of data (eg, infected and not infected).¹¹ Fernandez-Delgado et al¹² evaluated 179 classifiers from 17 families, and the random forest algorithms were the best, achieving 94% accuracy. Considering the semiautomated method (man-and-machine approach), human review can improve the specificity and precision of algorithm performance in daily practice.¹¹ Furthermore, the model used several features that included most NHSN criteria, which may increase the generalizability of the algorithm application.

Proxy variables are commonly used for surveillance applications. In one study, patients were screened using proxy indicators such as antimicrobial use, white blood cell counts, and fever. Antimicrobial use identified 95% of those at higher risk of infection.¹³ Most NHSN criteria are related to proxy variables such as fever, white blood cell count, laboratory culture results, and radiology results such as consolidation, cavitation, and infiltrates.³ A few other criteria may be related to a condition that can increase the risk for infection, for example, the use of devices or other invasive procedures. Most of the criteria for immunosuppressed patients (eg, leukemia, lymphoma, HIV positive, splenectomy, chemotherapy, and those on steroids) may be related to risk factors for infection.³ The ML model included a combination of proxy features and risk-related variables. This combination performed better than proxy or risk factors features alone, especially in terms of sensitivity and positive predictive value. Notably, the model including only risk variables performed better than proxy variables, indicating that patient comorbidities or severity and other risk factors for infection could be used to identify patients for possible HAI evaluation and prevention.

Traditional surveillance techniques must be patient based, which means seeking infections during a patient's stay, and screening a variety of data: laboratory results; antibiotic use; admission, discharge, and transfer data; radiology results; pathology databases, patient charts, healthcare worker notes, physical exam notes, vital signs, and invasive device use. Laboratory-based surveillance should not be used alone, primarily because of a lack of sensitivity.^{3,14} In this study, manual surveillance based on laboratory culture results affected HAI classification results. The semiautomated method greatly improved sensitivity. The semiautomated methodology was not dependent only on laboratory culture results for HAI classification; it surveyed data from all inpatient care, including vital signs and clinical notes. This aspect was important for improving performance not only for cases dependent on laboratory culture results (eg, BSI and UTI) but also for infections with criteria that do not depend on culture results (eg, respiratory infection and SSI).

When we compared both methods for those infections that classification is dependent on laboratory culture results, semi-automated surveillance outperformed manual search by 31.6%, demonstrating that ML can overcome the difficulties of traditional manual surveillance.

Traditional surveillance methods have limitations, and AI initiatives can overcome some of these drawbacks, such as fatigue, strenuous activities, and lack of time. Russo et al⁷ demonstrated that electronic surveillance maintained high levels of sensitivity

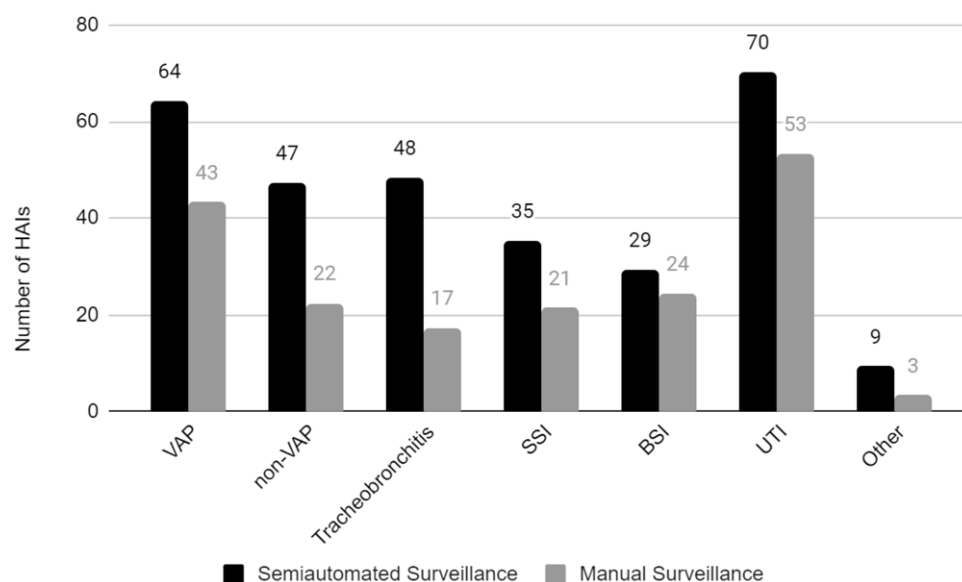


Figure 1. Comparison between semiautomated surveillance and manual surveillance regarding the number of infections identified per site. Note. HAIs, hospital-acquired infections; VAP, ventilator-associated pneumonia; SSI, surgical-site infection; BSI, bloodstream infection; UTI, urinary tract infection.

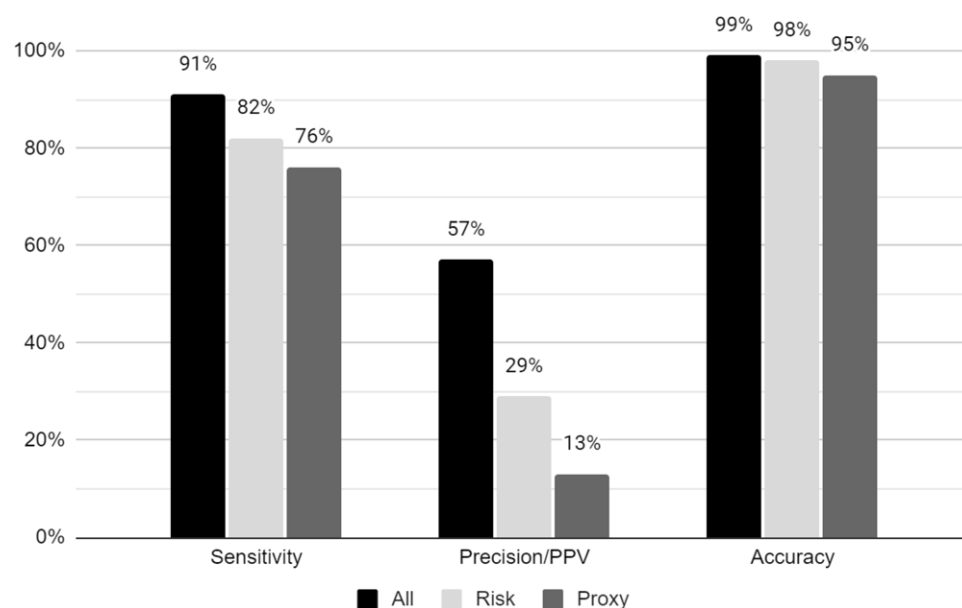


Figure 2. Model performance (sensitivity, precision/PPV, accuracy) considering all 447 features (all), risk variables (risk), and proxy variables (proxy).

(84%–100%) and specificity (88%–100%) and a reduction effect in time spent on infection prevention of 50%–90%. One of the applications of AI is augmenting human capabilities at the same time that it reduces workload.⁹ Scardoni *et al*⁸ reviewed 27 studies on AI tools for HAI surveillance and found moderate evidence that ML-based models perform equal to or better than non-ML approaches. One Swedish study reported a sensitivity of 93.7% and positive predictive value of 79.7% for retrospective HAI classification using gradient boost algorithms.¹⁷ In Geneva, researchers reported high sensitivity (88.6%–92.6%) using support-vector machine algorithms for nosocomial infection identification.^{18–20} The semiautomated surveillance method studied here has shown a sensitivity of 97% and specificity of 98.2%.⁹

In this single-center ML study, we used 1 database and 1 gold standard for performance comparisons. Performance results may change related to the reference golden standard chosen or different data analysis by the algorithm. However, our results are promising because the algorithm criteria overlap with the NHSN criteria and

other features and can search all hospital inpatients. Also, the use of automated systems can reduce human variability in criteria identification. ML algorithms based on clinical data in the EHR depend on the completeness of data for performance. A poor EHR may underestimate HAIs in both manual and automated surveillance. Using large amounts of data and hundreds of features may limit implementation because computational capacity requirements are greater. Enhancing the performance of algorithms and reducing the number of features are highly desirable objectives to facilitate the wider dissemination and adoption of this technology.

In conclusion, the ML algorithm, which included most NHSN criteria and >200 other features, augmented the human capacity for HAI classification. The combination of features, including proxy variables and risk factors for infections, achieved best performances. Well-documented and understandable algorithm performances may facilitate the use and incorporation of AI tools in clinical or epidemiological practice and can help overcome the drawbacks of traditional HAI surveillance.

Acknowledgments.

Financial support. No financial support was provided relevant to this paper.

Conflict of interest. Stephani Amanda Lukasewicz Ferreira, Arateus Crysham Franco Meneses, Tiago Andres Vas, Otavio Luiz da Fontoura Carvalho, Camila Hubner Dalmora, and Rodrigo Pires dos Santos are members of Qualis.

References

1. Magill SS, O'Leary E, Janelle SJ, *et al.* Changes in prevalence of healthcare-associated infections in US hospitals. *N Engl J Med* 2018; 379:1732–1744.
2. Weiner-Lastinger LM, Pattabiraman V, Konnor RY, *et al.* The impact of coronavirus disease 2019 (COVID-19) on healthcare-associated infections in 2020: a summary of data reported to the National Healthcare Safety Network. *Infect Control Hosp Epidemiol* 2022;43:12–25.
3. National Healthcare Safety Network. HAI checklists. Centers for Disease Control and Prevention website. <https://www.cdc.gov/nhsn/hai-checklists/index.html>. Accessed June 27, 2023.
4. Agência Nacional de Vigilância Sanitária. Critérios Diagnósticos de Infecções Relacionadas à Assistência à Saúde, 2017.
5. Mitchell BG, Hall L, Halton K, MacBeth D, Gardner A. Time spent by infection control professionals undertaking healthcare-associated infection surveillance: a multicentered cross-sectional study. *Infect Dis Health* 2016;21:36e40.
6. Shenoy ES, Branch-Elliman W. Automating surveillance for healthcare-associated infections: rationale and current realities (part I/III). *Antimicrob Steward Healthc Epidemiol* 2023;3:e25.
7. 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.
8. Scardoni A, Balzarini F, Signorelli C, Cabitza F, Odone A. Artificial intelligence-based tools to control healthcare associated infections: a systematic review of the literature. *J Infect Public Health* 2020;13:1061–1077.
9. Ferreira SAL, Meneses ACF, Vaz TA, *et al.* An effective infection surveillance assistant robot impacts care delivery by reducing burden on infection control professional staff. *N Engl J Catalyst* 2022;3(8):1–17.
10. Dos Santos RP, Silva D, Menezes A, *et al.* Automated healthcare-associated infection surveillance using an artificial intelligence algorithm. *Infect Prev Pract* 2021;3:100167.
11. Deo RC. Machine learning in medicine. *Circulation* 2015;132:1920–1930.
12. Fernández-Delgado M, Cernadas E, Barro S, Amorim D, Fernández-Delgado A. Do we need hundreds of classifiers to solve real-world classification problems? *J Mach Learn Res* 2014;15:3133–3181.
13. Magill SS, Hellinger W, Cohen J, *et al.* Prevalence of healthcare-associated infections in acute-care hospitals in Jacksonville, Florida. *Infect Control Hosp Epidemiol* 2012;33:283–291.
14. Glenister HM, Taylor LJ, Bartlett CL, Cooke EM, Sedgwick JA, Mackintosh CA. An evaluation of surveillance methods for detecting infections in hospital inpatients. *J Hosp Infect* 1993;23:229–242.
15. Wundavalli L, Agrawal US, Satpathy S, Debnath BR, Agnes TA. How much is adequate staffing for infection control? A deterministic approach through the lens of workload indicators of staffing need. *Am J Infect Control* 2020;48:609–614.
16. Bartles R, Dickson A, Babade O. A systematic approach to quantifying infection prevention staffing and coverage needs. *J Infect Control* 2018;46: 487–491.
17. Ehrentraut C, Ekholm M, Tanushi H, Tiedemann J, Dalianis H. Detecting hospital-acquired infections: a document classification approach using support vector machines and gradient tree boosting. *Health Informatics J* 2018;24:24–42.
18. Cohen G, Hilario M, Sax H, Hugonnet S, Pellegrini C, Geissbuhler A. An application of one-class support vector machine to nosocomial infection detection. *Stud Health Technol Inform* 2004;107:716–720.
19. Cohen G, Sax H, Geissbuhler A. Novelty detection using one-class Parzen density estimator. An application to surveillance of nosocomial infections. *Stud Health Technol Inform* 2008;136:21–26.
20. Cohen G, Hilario M, Sax H, Hugonnet S, Geissbuhler A. Learning from imbalanced data in surveillance of nosocomial infection. *Artif Intell Med* 2006;37:7–18.