Data use and effectiveness in electronic surveillance of healthcare associated infections in the 21st century: a systematic review

Jeroen S de Bruin, Walter Seeling, Christian Schuh

Editor's choice
Scan to access more
free content

Section for Medical Expert and Knowledge-Based Systems, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria

Correspondence to
Dr Jeroen S de Bruin, Section for Medical Expert and Knowledge-Based Systems, Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Medical University of Vienna, Spitalgasse 23, Vienna A-1090, Austria; jeroen.debruin@meduniwien.ac.at

ABSTRACT

Objective As more electronic health records have become available during the last decade, we aimed to uncover recent trends in use of electronically available patient data by electronic surveillance systems for healthcare associated infections (HAIs) and identify consequences for system effectiveness.

Methods A systematic review of published literature evaluating electronic HAI surveillance systems was performed. The PubMed service was used to retrieve publications between January 2001 and December 2011. Studies were included in the review if they accurately described what electronic data were used and if system effectiveness was evaluated using sensitivity, specificity, positive predictive value, or negative predictive value. Trends were identified by analyzing changes in the number and types of electronic data sources used.

Results 26 publications comprising discussions on 27 electronic systems met the eligibility criteria. Trend analysis showed that systems use an increasing number of data sources which are either medico-administrative or clinical and laboratory-based data. Trends on the use of individual types of electronic data confirmed the paramount role of microbiology data in HAI detection, but also showed increased use of biochemistry and pharmacy data, and the limited adoption of clinical data and physician narratives. System effectiveness assessments indicate that the use of heterogeneous data sources results in higher system sensitivity at the expense of specificity.

Conclusions Driven by the increased availability of electronic patient data, electronic HAI surveillance systems use more data, making systems more sensitive yet less specific, but also allow systems to be tailored to the needs of healthcare institutes' surveillance programs.

INTRODUCTION

Healthcare associated infections (HAIs) are generally recognized as a considerable threat to a patient's health, as they are associated with increased morbidity and mortality,[1–4] and increase the cost of healthcare.[5–7] To decrease the rate of infection, hospitals have adopted active surveillance programs which have proven to be effective,[8–10] but are also very time-demanding for infection control specialists,[11 12] and the results are susceptible to considerable variability.[13 14] Consequentially, electronic HAI surveillance systems have been developed to decrease variability and allow infection control specialists to focus less on infection detection and more on infection prevention.

To detect HAIs, electronic surveillance systems utilize electronically available patient data, such as clinical, microbiological, pharmaceutical, and administrative patient records. Over the last decade, more types of electronic health records have become available in hospitals, providing opportunities to improve the effectiveness of electronic HAI detection systems in the detection of both old and new threats. We initiated this systematic review to assess more recent trends in electronic data usage by HAI surveillance systems and resulting consequences for system effectiveness by analyzing systems created in the first decade of the 21st century.

METHODS

Search strategies and information sources

We conducted a systematic search of published literature that evaluated electronic surveillance systems for HAIs. Searches were conducted both electronically and manually; we used the PubMed service to search for publications indexed in Medline between January 1, 2001 and December 31, 2011; manual searches were performed by scanning the bibliographies of all eligible original research papers and systematic reviews, as well as the authors' personal collections. Figure 1 shows the PubMed search query. We used the filters 'human' and 'abstract' on all searches, and the publication language was restricted to English.

Eligibility criteria and study selection

Included articles had to describe a system that performs electronic HAI surveillance, what electronically available data sources were used by the system, and how. Reviews were excluded, as were publications addressing modifications to already published systems, to avoid overcounting data sources. There were no restrictions to monitored infection types, electronic data sources, or healthcare settings. To evaluate the effectiveness of a system, sensitivity, specificity, positive predictive value (PPV), or negative predictive value (NPV) either had to be stated or could be calculated from the raw data.

Titles and abstracts were evaluated independently by all authors to confirm an article matched the aforementioned profile. Discrepant cases were settled by a majority wins consensus. Relevant complete articles were retrieved and reviewed independently.

Data collection and extraction

Data were abstracted from each article to a standardized Excel worksheet. JSdB reviewed all articles, while ChS and WS each reviewed half and supervised the results. Collected data included patient characteristics, hospital setting, infection types

```
(((("2001/01/01"[Date - Publication] : "2011/12/31"[Date - Publication]))
AND
((((((healthcare-associated[Title/Abstract]) OR health care-
associated[Title/Abstract]) OR hospital-acquired[Title/Abstract]) OR hospital
acquired[Title/Abstract]) OR surgical site[Title/Abstract]) OR postoperative
wound[Title/Abstract]))
AND
((((((electronic*[Title/Abstract]) OR computer*[Title/Abstract]) OR
automate*[Title/Abstract]) OR registry[Title/Abstract]) OR database[Title/Abstract]))
AND
((((((surveillance[Title/Abstract]) OR monitor*[Title/Abstract]) OR
detect*[Title/Abstract]) OR report*[Title/Abstract]) OR prevent*[Title/Abstract]) OR
control*[Title/Abstract])
```

Figure 1 The PubMed search query for electronic healthcare-associated infection surveillance systems.

targeted by the system, types of electronic data sources used and how, type of algorithms used, and overall performance measures. In case multiple values for performance measures were reported, the value indicated as most favorable by the article authors was extracted; if an overall performance indication was not specified, we attempted to calculate it from the available data, or specified a range if recalculation proved impossible. For each article, data elements that were not confirmed by all reviewers were discussed until a consensus was reached.

Data analysis
Discussion of publications was divided into two periods; the first period contains all studies from 2001 to 2006 (period 1), whereas the second period spans 2007–2011 (period 2). Based on this partitioning, we analyzed changes in the number, type, and category of electronic data sources used, and compared overall system effectiveness. We also present graphical depictions of the aforementioned statistics in 5-year sliding windows between period 1 and period 2.

RESULTS
The aforementioned search strategy generated 431 results; 380 references were found in the electronic search, 20 in the authors' personal libraries, and 31 in bibliographies of reviews and eligible research articles. After duplicate removal and review of titles and abstracts, 85 articles were considered eligible for further study. Fifty-five full-text articles were found by the authors, and an additional 22 through the help of a senior librarian. Twenty-six publications satisfied the aforementioned eligibility criteria and were included in the review. Figure 2 shows the study flow diagram.

Electronic data types were categorized as either *medico-administrative* or *clinical- and laboratory-based* data. Medico-administrative data includes procedure and discharge coding, pharmacy dispensing records, and physician narratives, but excludes patient admission and tracking data. Clinical- and laboratory-based data includes data on patient vital signs, as well as biochemistry, microbiology, and radiology laboratory test results. Data categories are eponymous to system categories that use these data categories exclusively, while systems using data from both categories are called mixed-data systems.

The included systems were designed for a variety of HAIs, including surgical site infections (SSIs) and postoperative wound infections (PWIs),[15–26] urinary tract infections (UTIs),[16 20 22 25 27–30] bloodstream infections (BSIs),[20 22 27–29] central venous catheter-related infections (CRIs) and central line-associated BSIs (CLABSIs),[23 27–29 31–33] pneumonias and lower respiratory tract infections (RTIs),[20 22 23 28 29 34–36] *Clostridium difficile* infections (CDIs),[20 37] and drain-related meningitis.[38]

Period 1: 2001–2006
Period 1 comprised 13 electronic HAI detection systems. Table 1 provides an overview of all publications in period 1.
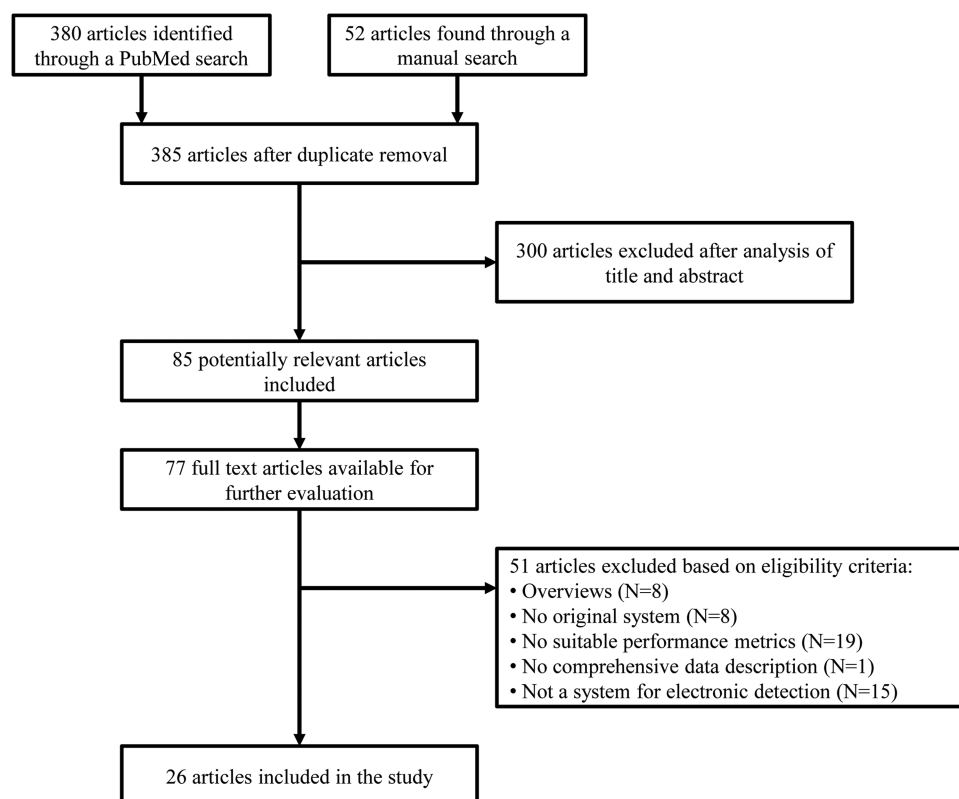
Figure 2 Study selection flow diagram.

Table 1 Electronic healthcare-associated infection surveillance systems published from 2001 to 2006

Article	Study setting and size	HAI types	Data sources used	Data description	Sensitivity [PPV] (%)	Specificity [NPV] (%)
Cadwallader et al[15]	Adult teaching hospital, 510 procedures	SSI	Procedure and discharge codes	ICD-9-CM codes	88 [95.6]	99.8
Yokoe et al 2001[16]	HMO, 2746 female patients	SSI, UTI, PWI	Procedure and discharge codes; pharmacy dispensing records	ICD-9-CM and, COSTAR codes; antimicrobial exposure	96 [40]	99
Bouam et al[27]	Teaching hospital, 548 samples	UTI, BSI, CRI	Microbiology laboratory results	Positive blood, urine or CVC cultures	91 [88]	91 [93]
Moro and Morsillo[17]	HMO, 6158 patients	SSI	Procedure and discharge codes	ICD-9-CM codes	20.6	99.1
Trick et al[31]	Teaching and community hospitals, 127 patients	CRI	Microbiology laboratory results; pharmacy dispensing records	Positive blood or wound cultures; vancomycin exposure interval	81 [62]	72 [87]
Yokoe et al[18]	HMO, 22 313procedures	SSI	Procedure and discharge codes; pharmacy dispensing records	ICD-9-CM codes; infection-specific antimicrobial exposure intervals	93–97 [33–38]	–
Spolaore et al[19]	Acute care hospitals, 865 cases	SSI	Procedure and discharge codes; microbiology laboratory results	ICD-9-CM codes; positive wound and drainage cultures	[97]	–
Mendonça et al[34]	Acute care and community hospitals, 1688 neonates	Pneumonia	Radiology results	Free text radiology reports	71 [7.5]	99
Leth and Møller[22]	Teaching hospital, 1129 patients	UTI, BSI, pneumonia, SSI	Pharmacy dispensing records; microbiology laboratory results; biochemistry laboratory results	Antimicrobial exposure; positive blood, urine, wound and drainage cultures; abnormal leukocyte counts and CRP values	94 [21]	47 [98]
Brossette et al[20]	Tertiary teaching and community hospitals, 907 patients	UTI, BSI, RTI, CDI, PWI	Microbiology laboratory results	Positive blood, urine, sputum and wound cultures, serology, molecular tests	86	98.4
Sherman et al[23]	Academic tertiary care pediatric hospital, 1072 cases	CRI, VAP, SSI	Procedure and discharge codes	ICD-9-CM codes	61 [20]	96 [99]
Pokorny et al[39]	Acute care teaching hospital, 1043 patients	Not specified	Procedure and discharge codes; pharmacy dispensing records; microbiology laboratory results	ICD-9-CM codes; antimicrobial exposure; positive cultures	94.3 [55.9]	83.6 [98.5]
Chalfine et al[21]	Tertiary care hospital, 766 patients	SSI	Microbiology laboratory results	Positive cultures	84.3 [96.4]	99.9 [99.4]

BSI, bloodstream infection; CDI, Clostridium difficile-related infection; COSTAR, computerized stored ambulatory records; CRI, CVC-related infection; CRP, C reactive protein; CVC, central venous catheter; HAI, healthcare-associated infection; HMO, health maintenance organization; ICD-9-CM, International Classification of Diseases 9th revision, clinical modification; NPV, negative predictive value; PPV, positive predictive value; PWI, postoperative wound infection; RTI, respiratory tract infection; SSI, surgical site infection; UTI, urinary tract infection; VAP, ventilator-associated pneumonia.

Clinical- and laboratory-based systems

Bouam et al[27] discussed a knowledge-based approach to identify a variety of HAIs using a series of computer programs to extract positive blood, urine, or catheter culture results. The system achieved a sensitivity and specificity of 91%, outperforming the concurrently performed ward-based surveillance. A similar search strategy including serology and molecular test results as well as a rule base for the exclusion of samples likely associated with specimen contamination and non-infected clinical states was discussed by Brossette et al.[20] The system missed seven infections (sensitivity 86%) and generated 14 false positives (specificity 98.4%). Chalfine et al[21] discussed a combination of automatic positive culture recognition for surgical site specimens, and manual confirmation of results by surgeons. Despite the manual confirmation, the sensitivity (84.3%) was not higher than the aforementioned automated approaches. Mendonça et al[34] reported on a natural language processing system that analyses free-text radiology results for infection indications to detect hospital-acquired pneumonias among neonates. The resulting system was very specific (99.9%), and showed decent sensitivity (71%), but very poor PPV (7.5%) due to 61 false positives.

Medico-administrative-based systems

Cadwallader et al[15] discussed an SSI detection method that searches patient records for selected procedure- and infection-specific International Classification of Diseases 9th revision, clinical modification (ICD-9-CM) discharge codes, which yielded high sensitivity (88%), specificity (99.8%), and PPV (95.6%). However, both Moro and Morsillo[17] and Sherman et al[23] report lower sensitivities (respectively, 62% and 20.6%) with similar search strategies, despite using more comprehensive collections of ICD-9-CM codes.

Yokoe et al[16] describe a method which used procedure and discharge codes and pharmacy records for selected drug types as input for univariate data analysis to select postpartum infection predictors. These predictors served as input for a predictive logistic regression model, which achieved 96% sensitivity and 99% specificity. A rule-based system analyzing the individual and collective effectiveness of procedure-specific intervals of antimicrobial drug treatment and ICD-9-CM codes to detect a variety of SSIs was described in Yokoe et al,[18] and yielded sensitivities of 93–97%, depending on the infection site.

Mixed-data systems

Trick et al[31] described a rule-based classification system which uses positive culture results and vancomycin exposure to detect central venous CRIs. The system sequentially classifies potential infections as either hospital- or community-acquired, infection or contamination, and as primary or secondary infections, excluding cases with each failed classification. The automated system yielded good sensitivity (81%) and decent specificity (72%); a semi-automated version with manual catheter determination increased specificity to 90%. Spolaore et al[9] analyzed the effectiveness of infection-associated ICD-9-CM discharge codes and positive wound and drainage cultures for tracking suspected SSIs. Where detection using individual data types was moderately effective (PPV: 70% for both types), the combination yielded excellent PPV (97%). A similar study on the effectiveness of positive microbiology cultures, recorded antibiotic therapy, and ICD-9-CM diagnostic codes was performed by Pokorny et al.[39] The authors found that a combination of two criteria resulted in the most satisfactory sensitivity (94.3%) and

specificity (83.6%). Finally, Leth and Møller[22] described a system that used biochemistry test results for leukocyte counts and C reactive protein (CRP) levels, positive blood, urine, wound, or drainage cultures, and recorded antimicrobial drug administration to identify a variety of HAIs. The combination of all data types proved most sensitive (94%), but not very specific (47%); when also considering community-acquired infections, the system showed good sensitivity, specificity, PPV, and NPV for all infection sites.[22]

Period 2: 2007–2011

Period 2 comprises 14 electronic HAI detection systems discussed in 13 publications. An overview of all publications in period 2 is shown in table 2.

Clinical- and laboratory-based systems

Bellini et al described a multi-phase classification algorithm similar to that of Trick et al using positive blood and catheter cultures to classify CRIs and BSIs, and discussed improvement strategies on pathogen inclusion. After combining data for both CRIs and BSIs, we calculated overall system sensitivity at 89.7%.[33] Woeltje et al discussed a rule-based system that was meant to augment manual surveillance of CLABSIs, and therefore optimized the system's NPV. Various rules were constructed and evaluated, some including antimicrobial treatment and patient vital signs, but a rule using only positive blood culture results yielded the most satisfactory NPV (99.2%) and specificity (68%).[32]

Klompas et al[35] described a system to identify ventilator-associated pneumonia cases in intensive care units (ICUs). Suspected cases are systematically excluded by analyzing chest radiograph results, abnormal leukocyte counts, the presence of fever, Gram stain results of respiratory secretion samples, and sustained increases in ventilator gas exchange. Although a true gold standard was lacking, all identified suspected cases were confirmed by infection control experts (sensitivity: 95%; PPV: 100%). Another ICU-based system is described by Koller et al.[28] The system combined positive cultures, abnormal leukocyte counts and CRP levels, and patient vital signs in a knowledge base using fuzzy sets and logic. The system showed good outcomes for all performance metrics, having missed only three of 21 cases and reporting no false positives. Finally, Choudhuri et al discussed a multi-phase rule-based classification system for the detection of catheter-associated UTIs using positive microbiology results, the presence of fever and catheters, and urinalysis results. The system showed good sensitivity (86.4%) and specificity (93.8%).[30]

Medico-administrative-based systems

Shaklee et al discussed a system to automatically detect CDI in children. While trying a variety of data sources including microbiology and antibiotic treatment records, the best sensitivity was achieved using only ICD-9-CM codes (sensitivity 80.7%; specificity 99.9%).[37] Bouzbid et al reviewed two different methods for detecting a variety of HAIs, one using natural language processing on electronic hospital discharge summaries created by a medical epidemiologist through the manual data extraction of patient vital signs, antibiotic or antifungal prescriptions, and positive microbiology reports. Though only used on a small subset of the study population, the method showed the highest sensitivity/specificity ratio (sensitivity 86.7%; specificity 88.2%).[29]

Bolon et al reviewed the effectiveness of rules using ICD-9-CM codes and infection-specific antimicrobial exposure during index and rehospitalization to track infections after arthroplastic hip and knee procedures. They found that

Table 2 Electronic healthcare-associated infection surveillance systems published from 2007 to 2011

Article	Study setting and size	HAI types	Data sources used	Data description	Sensitivity [PPV] (%)	Specificity [NPV] (%)
Bellini et al[33]	University hospital, 669 episodes	BSI, CRI	Microbiology results	Positive blood and CVC cultures	89.7*	83.9*
Klompas et al[35]	Academic hospital, 459 patients	VAP	Microbiology results; biochemistry results; radiology results; clinical patient data	Gram stains of excretion samples; abnormal leukocyte count; radiological signs of pneumonia; PEEP and FiO_2 values, presence of fever	95 [100]	–
Woeltje et al[32]	Tertiary care academic hospital, 540 patients	CRI	Microbiology results	Positive wound, urine or respiratory device cultures	94.3 [22.8]	68 [99.2]
Claridge et al[36]	Level I trauma center, 769 patients	VAP	Pharmacy dispensing records; microbiology results; clinical patient data	Antimicrobial treatment records, positive cultures, patient vital signs and presence of devices	97 [100]	100 [99.9]
Bolon et al[24]	HMO, 6322 procedures	SSI	Procedure and discharge codes; pharmacy dispensing records	ICD-9-CM codes; Antimicrobial administration with an infection-specific interval	86–93 [25–39]	–
Leth et al[25]	Community hospitals, 1512 women	SSI, UTI	Procedure and discharge codes; pharmacy dispensing records; microbiology results	ICD-10 and NCSP codes; ATC codes for antimicrobial administration; positive urine or wound cultures	74* [81.7*]	99.4* [99.3*]
Koller et al[28]	Tertiary care and teaching hospital, 99 patients	UTI, BSI, CRI, pneumonia	Microbiology results; biochemistry results; clinical patient data	Positive cultures; Abnormal leukocyte count or CRP values; Clinical data from a patient data management system	90.3 [100]	100 [95.8]
Shaklee et al[37]	Pediatric hospitals, 119 patients	CDI	Procedure and discharge codes	ICD-9-CM codes	80.7 [73.95]	99.9 [99.9]
Inacio et al[26]	HMO, 42 173 procedures	SSI	Procedure and discharge codes; physician narratives	ICD-9-CM codes; standardized postoperative forms	97.8 [11]	91.5 [100]
Bouzbid et al[29]	University hospital, 1499 patients	UTI, BSI, CRI pneumonia	System 1 physician narratives	Electronic discharge summaries	86.7	88.2
			System 2 pharmacy dispensing records; microbiology results	ATC codes for antimicrobial administration; positive cultures of non-common skin contaminants	99.3 [34.7]	56.8 [99.7]
Choudhuri et al[30]	University affiliated urban teaching hospital, 136 patients	Catheter-associated UTI	Microbiology results; biochemistry results; clinical patient data	Positive cultures; presence of fever and urinary tracts; abnormal leukocyte count	86.4 [85]	93.8 [94.4]
van Mourik et al[38]	Tertiary healthcare centre, 537 patients	Drain-related meningitis	Pharmacy dispensing records; microbiology results; biochemistry results	Number and exposure time of antimicrobial drugs; positive cerebrospinal fluid and drain cultures; abnormal leukocyte and CRP values	98.8 [56.9]	87.9 [99.9]
Chang et al[40]	Academic teaching medical center, 476 patients	Device-related	Procedure and discharge codes; pharmacy dispensing records	ICD-9-CM codes; administration of steroids	92.4–96.6 [70.6–79.1]	86.0–91.5 [97.2–98.7]

*Indicates that metrics were (re-)calculated from data provided in the publication.
ATC, Anatomical Therapeutic Chemical; BSI, bloodstream infection; CRI, CVC-related infection; CRP, C reactive protein; CVC, central venous catheter; HAI, healthcare-associated infection; HMO, health maintenance organization; ICD-9-CM, International Classification of Diseases 9th revision, clinical modification; NCSP, NOMESCO Classification of Surgical Procedures; NPV, negative predictive value; PEEP, positive end-expiratory pressure; PPV, positive predictive value; SSI, surgical site infection; UTI, urinary tract infection; VAP, ventilator-associated pneumonia.

combining both data sources resulted in high sensitivity (86–93%) but poor PPV (25–39%) for both infection sites.[24] Chang et al also used combinations of ICD-9-CM codes and steroid and chemo treatment indications as inputs for a predictive logistic regression model to uncover parameter sets significant to device-related HAIs. These parameter sets also served as inputs to an artificial neural network. Using all available parameters, the neural network proved more sensitive (96.6%), while the regression model was more specific (91.5%).[40] A semi-automated method requiring a manual confirmation step to detect SSIs was introduced by Inacio et al.[26] Electronic health records were scanned for relevant ICD-9-CM codes, while simultaneously standardized surgical reports pertaining information on all surgical procedures were searched for infection-related terms. Positive outcomes for either search strategy were combined and manually verified by infection control experts, which achieved excellent sensitivity (97.8%) and high specificity (91.5%).

Mixed-data systems

The second method discussed in Bouzbid et al analyzed the effectiveness of positive blood cultures, anatomical therapeutic chemical (ATC) codes for systemic antibiotic exposure, and ICD-10 codes. After an exhaustive review of all combinations, using either microbiology or antibiotic prescription produced the best sensitivity (99.3%), at the expense of specificity and PPV (respectively, 56.8% and 34.7%).[29] Leth et al[25] described a system to detect in-hospital and post-discharge PWIs and UTIs using sets of procedure and discharge codes, infection-specific ATC codes, and positive urine or wound cultures. After combining data from all contingency tables, the system showed overall sensitivity of 74% and specificity of 99.4%, although it is worth noting that in-house infection detection was more effective than post-discharge detection.[25] Claridge et al evaluated a system that detects ventilator-associated pneumonia in ICU patients using positive microbiology results, patient vital signs and recorded use of ventilators, and data on current antibiotic treatments. The resulting combination was highly accurate

(sensitivity 97%; specificity 100%).[36] Finally, van Mourik et al analyzed the effectiveness of a logistic regression model for drain-related meningitis, thereby using positive drain cultures, abnormal biochemical blood and drain-related values, and the number and exposure intervals of antibiotics as input. The resulting predictive model yielded excellent sensitivity (98.8%) and high specificity (87.9%).[38]

Data trends

The average number of data sources used by electronic HAI detection systems rose from 1.62 per system in period 1 to 2.21 in period 2. The 5-year trend movement (figure 3) showed that until recently, systems used clinical- and laboratory-based data more often than medico-administrative-based data. The figure also shows that about one-third of all systems between period 1 and period 2 depended on both data types.

Trends in the use of different data types (figure 4) showed that microbiology data use has increased over the last decade and is the most frequently used data type in electronic HAI detection systems; a modest increase in the use of pharmacy data and a substantial increase in the use of biochemistry data were also detected. Period 2 also saw the introduction of systems using clinical data such as patient vital signs, and more recently modest attempts at the use of electronic physician narratives. In contrast, substantially fewer systems have used diagnostic and treatment codes over the last decade; approximately one-third of all systems use this type of data. Systems using radiology data remain a rare occurrence, though the frequency of use remained unchanged.

Figure 5 shows a comparison between the two periods for all performance metrics. For sensitivity, PPV, and NPV, results are generally better in period 2, while systems in period 1 achieve better specificity. The figure also shows that for all except PPV, variability between systems in period 2 is lower. The receiver operating characteristic graph in figure 6 shows the sensitivity/specificity ratio for all systems that reported on sensitivity and specificity for both periods. Systems for which multiple performance metrics were reported are depicted as ranges. The

Figure 3 Distribution of systems using only clinical- and laboratory-based data, medico-administrative data, or both (mixed data) in 5-year sliding periods between 2001 and 2011.
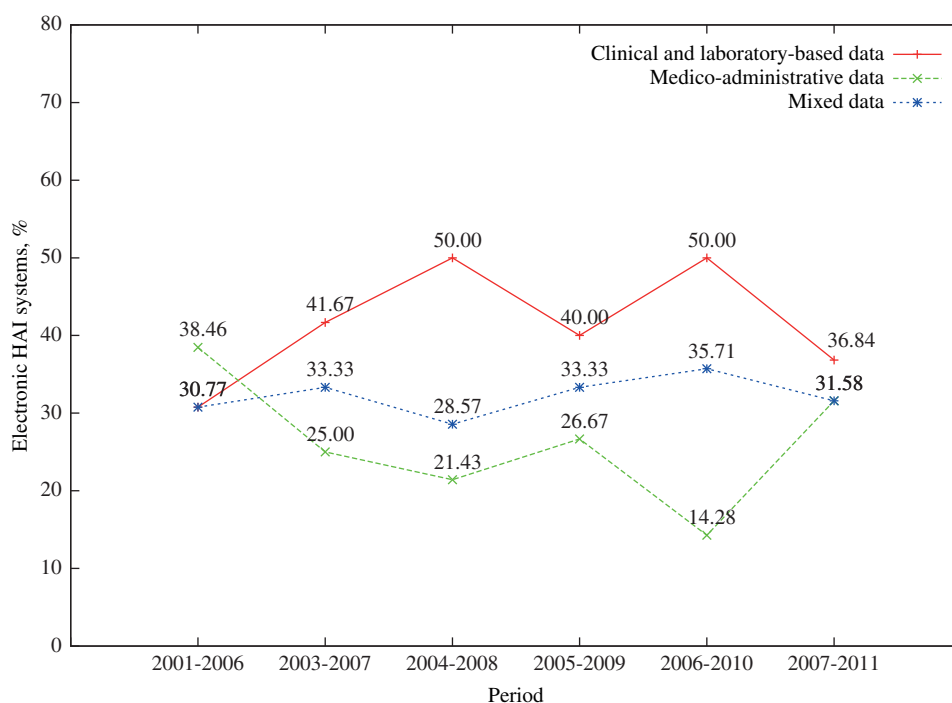
Figure 4 Frequency of use for each data source type in 5-year periods between 2001 and 2011.
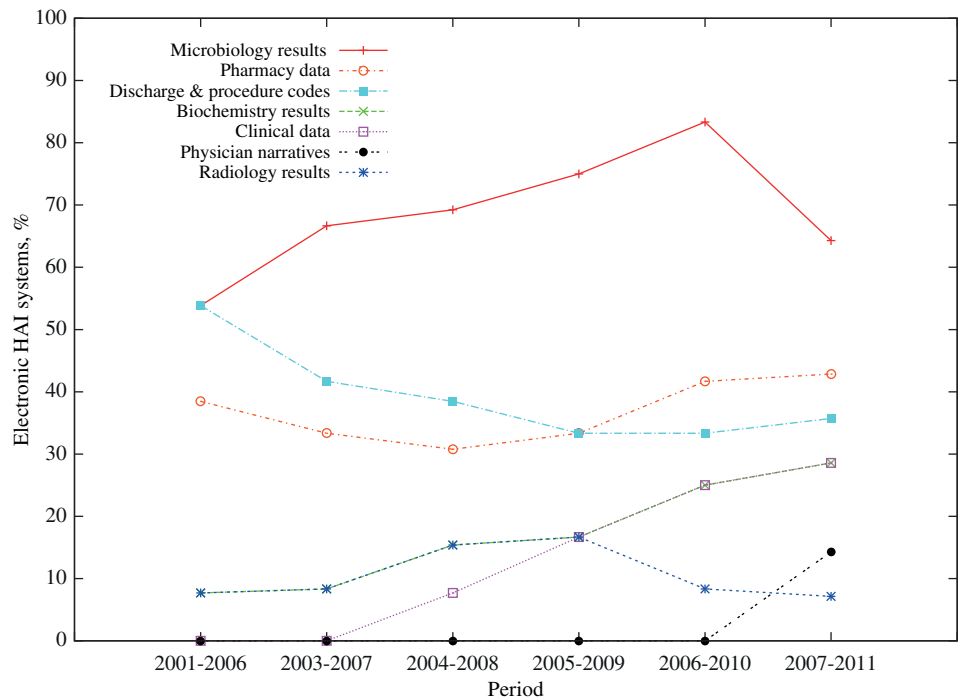
graph shows that more recent systems sacrifice specificity to achieve a high sensitivity.

DISCUSSION

This review showed the developments in availability and use of electronic data in the electronic detection of HAIs during the first decade of the 21st century, and how these developments influenced detection effectiveness. The benefits of electronic surveillance have often been addressed; some studies indicated that surveillance rules are complex and open to subjective interpretation,[27 35] causing considerable variability in manual surveillance results, both within and between healthcare institutes.[24 34 41] Electronic surveillance systems lack variability as they consistently apply surveillance definitions[23 24 31 35]; studies that directly compared electronic and manual surveillance performance showed that electronic surveillance achieves equal or better sensitivity than manual surveillance.[11 15 18 20 24 27 31 36] Several studies also reported time savings of 60–99.9%,[20 21 27 28] or a reduction in chart reviews of 40–90.5%.[26 29 32 38]

Despite the reported successes of electronic surveillance, only eight systems (25%) were used in clinical routine.[20 21 26–28 30 34 36] A survey among the authors (response rate 67%) revealed an additional two systems used in clinical routine.[11 32] Reasons mentioned for not adopting methods in clinical or organizational routine were

Figure 5 Per-metric performance comparisons between systems published in 2001–2006 (period 1) and in 2007–2011 (period 2). For each performance metric, period 1 is plotted on the left, and period 2 on the right. PPV, positive predictive value; NPV, negative predictive value.
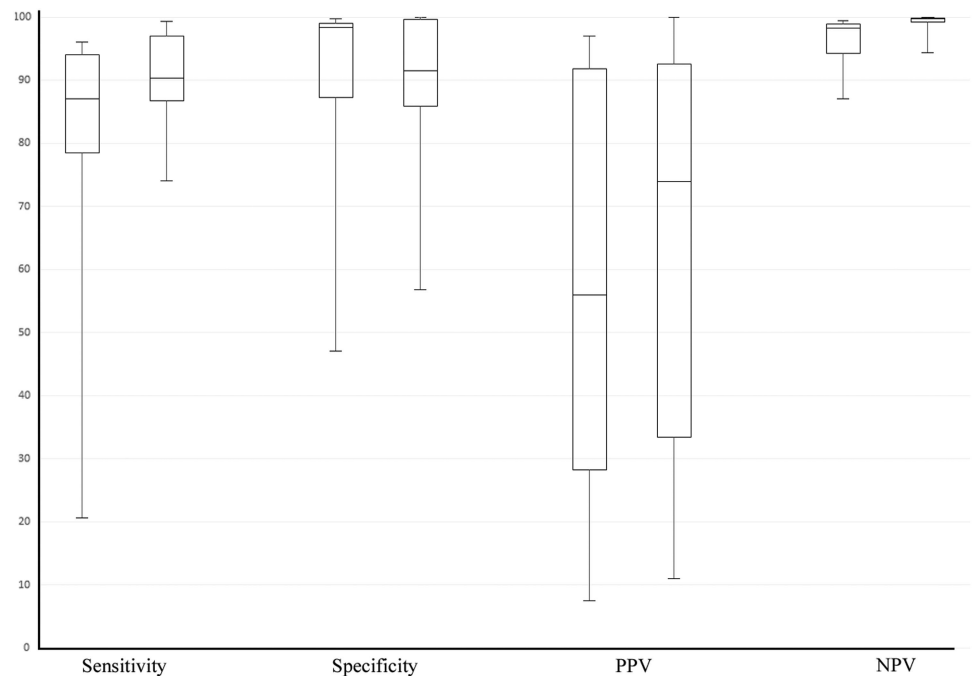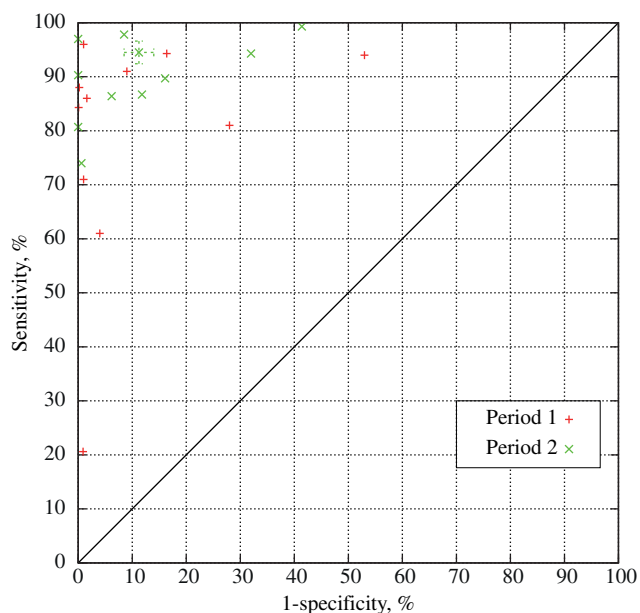
Figure 6 The receiver operating characteristic graph for systems published in 2001–2006 (period 1) and in 2007–2011 (period 2).

a lack of resources to manage technical challenges, unfavorable results of the system, and that systems were used for research only.

The increase in the number of data sources used by electronic HAI detection systems is driven by an overall increase in availability of electronic health records and the need for more data sources to increase system accuracy.[42] Several systems reported on potential improvements if more patient data were electronically available[27 31 33 34]; other studies found combining data sources can considerably improve the effectiveness of a system.[18 22 29 39]

The frequent use of microbiology data can be explained by its high electronic availability in hospitals and its clinical importance and effectiveness in predicting HAIs.[32] Several studies suggest that the percentage of culture-negative infections is 5–20%, depending on infection site,[38 43 44] which corresponds well with the sensitivities of reviewed systems based on microbiology data.[20 21 27 33 39] The effectiveness of microbiology data in surveillance depends on the proportion of suspected HAIs for which a microbiology culture is performed,[21] which limits its use in SSI and PWI surveillance where a relatively small number of specimens are evaluated due to a patient's short postoperative length of stay.[18 45] This limitation is reflected in this study; out of 12 systems that detect SSIs for various infection sites, only five used microbiology data (42%).

Studies on the effectiveness of diagnostic and procedure code data in HAI detection produced inconclusive results, which could explain its decline in use over the last decade. Multiple studies showed that diagnostic codes data was a weak indicator for HAIs.[17 23 46 47] Sherman et al provide several reasons for the decreased performance, stating that codes were not assigned by clinicians and therefore do not make use of all available clinical data, and that codes were not designed to differentiate HAIs from community-acquired infections.[23] Several studies also found high variability and inaccuracy in the use of billing codes both within and between healthcare institutes.[19 24 46] More recent studies stated that the use of ICD-9-CM codes *can* result in high sensitivity and specificity rates, and that limitations in sensitivity are likely caused using a small number of codes.[24 26 37]

HAI detection using antimicrobial drug administration records results in great sensitivity, but only moderate specificity.[18 22 24 29 48]

Antimicrobial drugs are often used for purposes other than infection treatment, such as postoperative prophylaxis, and for some procedures the patient population is too heterogenic to be an effective predictor by itself.[21 32 39] Systems that used pharmacy data in combination with other data sources have shown excellent sensitivity while retaining a reasonable specificity by addressing the aforementioned problems in a variety of ways, such as introducing infection-specific filters on the number, types, or exposure of antibiotics used.[18 22 25 32 36 38] However, screening strategies based on pharmacy data need to be evaluated periodically because of changing antimicrobial subscribing practices.[24]

Several of the more recently introduced systems have used biochemical indicators of infection such as leukopenia, leucocytosis, and elevated levels of CRP,[22 38] patient vital signs such as the presence of fever or a rise in ventilator gas exchange,[32] or both.[28 30 35] A possible explanation could be an increased electronic availability of these types of data combined with their clinical importance in the detection of HAIs, as many of these signs have been linked to HAIs and adopted in surveillance programs.[49 50] Two studies indicated that the rules or knowledge base were designed to closely resemble established HAI detection rules.[28 35] However, systems that used patient vital signs were limited to the ICU setting due to the frequent usage of clinical monitoring systems and devices in this setting.[28 32 35]

Radiology data is rarely used because of its limited applicability and low electronic availability. In established surveillance programs, radiology results are primarily used for the detection of pneumonia.[49 50] We reviewed six systems that were designed to detect hospital-acquired pneumonia, but only two used electronically available radiology data[34 35]; Leth and Møller[22] also discussed the use and effectiveness of radiology data in predicting pneumonia, but radiology results were not used in the HAI-specific prediction model. The electronic format of radiology reports can also hinder its application; the analysis of electronically available free-text reports is a difficult task, which requires a supporting framework in order to be used effectively.[34]

Using physician narratives is the most recent trend in data usage, though several other studies already indicated its potential for electronic surveillance improvement.[22 25 31] The effectiveness of documentation for infection detection depends on its comprehensiveness. Inacio et al[26] used a relatively simple search strategy on surgical reports containing operation details, which identified only five additional cases of HAI that had not been identified by the ICD-9-CM search strategy. In contrast, Bouzbid et al used more comprehensive discharge summaries containing clinical and pharmacy data, as well as microbiology laboratory results, and the resulting system showed the best sensibility/specificity ratio compared to other approaches taken.[29] The type of documentation used influences system effectiveness as well; according to a recent study, discharge summaries contain the most information on infections missed by electronic surveillance, and electronic SSI detection would have the most benefit of using treatment reports.[51]

Due to the heterogenic study settings, patient populations, targeted infections, and performance indication of studies discussed in this review it was not possible to make a uniform comparison of all systems. However, figure 5 showed that in general, more recent systems achieved higher sensitivity and NPV but lower specificity; all performance metrics showed less variation for systems in period 2 as well. When analyzing the relationship between sensitivity and specificity for systems, figure 6 showed that more recent electronic HAI detection systems show a bias towards a higher sensitivity at the expense of specificity; several studies in period 2 also explicitly state a

preference for a higher sensitivity at the expense of specificity, or report on the method configuration which yields the highest sensitivity.[26 29 38]

This study has several limitations. First, the search was only done with PubMed, which may have limited the number of possible suitable publications; the review was based on 26 publications comprising discussions on 27 systems, which might be considered a low number compared to other reviews. Furthermore, eight publications were published in journals and proceedings for which no subscription was available and could not be obtained through other institutes; we are aware this could have biased the trends that we described. Finally, although a number of variations in search term have been used for completeness, publications focusing on specific types of HAIs may have been missed. For this reason, we added wound infection to the search terms, since very little publications were initially found on this subject. Consequentially, a bias towards SSIs and PWIs could have been introduced.

Several other reviews focused on electronic HAI surveillance systems. Leal and Laupland[52] wrote a review on the validity of electronic HAI detection systems compared to manual surveillance techniques. The review discussed 24 articles and used the same system categories as we did, but did not make a distinction between systems using one or multiple data sources. A very recent review by Freeman et al[53] also reports on advances made in electronic HAI surveillance in the 21st century and consequently shares many of the studies we evaluated. The review discussed 44 studies and focused on the utility of electronic HAI surveillance systems, thereby analyzing system effectiveness in the detection of different infection types. Though the review discussed more studies, several of their reviewed systems are in fact duplicates, and the review only differentiates between systems that do and do not use microbiology results. In this review we took a more comprehensive data-oriented approach that differentiated systems both in the number and types of data sources used for electronic detection. Furthermore, we focused on how changes in data use affected the performance of systems over the last decade, and provided detailed information on what information was used from individual data sources.

In conclusion, electronic HAI detection systems use increasingly more electronic health records and patient data as more data sources become available. As a result, systems tend to become more sensitive but less specific, though the increased availability allows systems to be configured and adapted in such a way that it can claim its own niche within a healthcare institute's surveillance program, leaving more time for infection control specialists to spend on infection prevention, and less on infection detection.

REFERENCES

1 European Centre for Disease Prevention and Control. Healthcare-associated infections. In Annual Epidemiological Report on Communicable Diseases in Europe 2008. Stockholm, Sweden; 2008;16–38. http://www.ecdc.europa.eu/en/publications/publications/0812_sur_annual_epidemiological_report_2008.pdf (accessed 10/06/2013).

2 Klevens RM, Edwards JR, Richards CL Jr, et al. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. Public Health Rep 2007;122:160–6.

3 Eggimann P, Pittet D. Infection control in the ICU. Chest 2001;120:2059–93.

4 Chen YY, Wang FD, Liu CY, et al. Incidence rate and variable cost of nosocomial infections in different types of intensive care units. Infect Control Hosp Epidemiol 2009;30:39–46.

5 Elward AM, Hollenbeak CS, Warren DK, et al. Attributable cost of nosocomial primary bloodstream infection in pediatric intensive care unit patients. Pediatrics 2005;115:868–72.

6 Sheng WH, Wang JT, Lu DC, et al. Comparative impact of hospital-acquired infections on medical costs, length of hospital stay and outcome between community hospitals and medical centres. J Hosp Infect 2005;59:205–14.

7 Warren DK, Quadir WW, Hollenbeak CS, et al. Attributable cost of catheter-associated bloodstream infections among intensive care patients in a nonteaching hospital. Crit Care Med 2006;34:2084–9.

8 Delgado-Rodriguez M, Gomez-Ortega A, Sillero-Arenas M, et al. Efficacy of surveillance in nosocomial infection control in a surgical service. Am J Infect Control 2001;29:289–94.

9 Burke JP. Surveillance, reporting, automation, and interventional epidemiology. Infect Control Hosp Epidemiol 2003;24:10–2.

10 Gastmeier P, Geffers C, Brandt C, et al. Effectiveness of a nationwide nosocomial infection surveillance system for reducing nosocomial infections. J Hosp Infect 2006;64:16–22.

11 Klompas M, Yokoe DS. Automated surveillance of health care-associated infections. Clin Infect Dis 2009;48:1268–75.

12 Wright MO, Perencevich EN, Novak C, et al. Preliminary assessment of an automated surveillance system for infection control. Infect Control Hosp Epidemiol 2004;25:325–32.

13 Emori TG, Edwards JR, Culver DH, et al. Accuracy of reporting nosocomial infections in intensive-care-unit patients to the National Nosocomial Infections Surveillance System: a pilot study. Infect Control Hosp Epidemiol 1998;19:308–16.

14 Glenister HM, Taylor LJ, Bartlett CL, et al. An evaluation of surveillance methods for detecting infections in hospital inpatients. J Hosp Infect 1993;23:229–42.

15 Cadwallader HL, Toohey M, Linton S, et al. A comparison of two methods for identifying surgical site infections following orthopaedic surgery. J Hosp Infect 2001;48:261–6.

16 Yokoe DS, Christiansen CL, Johnson R, et al. Epidemiology of and surveillance for postpartum infections. Emerg Infect Dis 2001;7:837–41.

17 Moro ML, Morsillo F. Can hospital discharge diagnoses be used for surveillance of surgical-site infections? Journal of Hospital Infection 2004;56:239–41.

18 Yokoe DS, Noskin GA, Cunnigham SM, et al. Enhanced identification of postoperative infections among inpatients. Emerg Infect Dis 2004;10:1924–30.

19 Spolaore P, Pellizzer G, Fedeli U, et al. Linkage of microbiology reports and hospital discharge diagnoses for surveillance of surgical site infections. J Hosp Infect 2005;60:317–20.

20 Brossette SE, Hacek DM, Gavin PJ, et al. A laboratory-based, hospital-wide, electronic marker for nosocomial infection: the future of infection control surveillance? Am J Clin Pathol 2006;125:34–9.

21 Chalfine A, Cauet D, Lin WC, et al. Highly sensitive and efficient computer-assisted system for routine surveillance for surgical site infection. Infect Control Hosp Epidemiol 2006;27:794–801.

22 Leth RA, Møller JK. Surveillance of hospital-acquired infections based on electronic hospital registries. J Hosp Infect 2006;62:71–9.

23 Sherman ER, Heydon KH, John KHS, et al. Administrative data fail to accurately identify cases of healthcare-associated infection. Infect Cont Hosp Ep 2006;27:332–7.

24 Bolon MK, Hooper D, Stevenson KB, et al. Improved surveillance for surgical site infections after orthopedic implantation procedures: extending applications of automated data. Clin Infect Dis 2009;48:1223–9.

25 Leth RA, Norgaard M, Uldbjerg N, et al. Surveillance of selected post-caesarean infections based on electronic registries: validation study including post-discharge infections. J Hosp Infect 2010;75:200–4.

26 Inacio MC, Paxton EW, Chen Y, et al. Leveraging electronic medical records for surveillance of surgical site infection in a total joint replacement population. Infect Control Hosp Epidemiol 2011;32:351–9.

27 Bouam S, Girou E, Brun-Buisson C, et al. An intranet-based automated system for the surveillance of nosocomial infections: prospective validation compared with physicians' self-reports. Infect Control Hosp Epidemiol 2003;24:51–5.

28 Koller W, Blacky A, Bauer C, et al. Electronic surveillance of healthcare-associated infections with MONI-ICU--a clinical breakthrough compared to conventional surveillance systems. Stud Health Technol Inform 2010;160(Pt 1):432–6.

29 Bouzbid S, Gicquel Q, Gerbier S, et al. Automated detection of nosocomial infections: evaluation of different strategies in an intensive care unit 2000–2006. J Hosp Infect 2011;79:38–43.

30 Choudhuri JA, Pergamit RF, Chan JD, et al. An electronic catheter-associated urinary tract infection surveillance tool. Infect Control Hosp Epidemiol 2011;32:757–62.

31 Trick WE, Zagorski BM, Tokars JI, *et al*. Computer algorithms to detect bloodstream infections. *Emerg Infect Dis* 2004;10:1612–20.

32 Woeltje KF, Butler AM, Goris AJ, *et al*. Automated surveillance for central line-associated bloodstream infection in intensive care units. *Infect Cont Hosp Ep* 2008;29:842–6.

33 Bellini C, Petignat C, Francioli P, *et al*. Comparison of automated strategies for surveillance of nosocomial bacteremia. *Infect Control Hosp Epidemiol* 2007;28:1030–5.

34 Mendonça EA, Haas J, Shagina L, *et al*. Extracting information on pneumonia in infants using natural language processing of radiology reports. *J Biomed Inform* 2005;38:314–21.

35 Klompas M, Kleinman K, Platt R. Development of an algorithm for surveillance of ventilator-associated pneumonia with electronic data and comparison of algorithm results with clinician diagnoses. *Infect Cont Hosp Ep* 2008;29:31–7.

36 Claridge JA, Golob JF, Fadlalla AM, *et al*. Who is monitoring your infections: shouldn't you be? *Surg Infect (Larchmt)* 2009;10:59–64.

37 Shaklee J, Zerr DM, Elward A, *et al*. Improving surveillance for pediatric Clostridium difficile infection: derivation and validation of an accurate case-finding tool. *Pediatr Infect Dis J* 2011;30:e38–40.

38 van Mourik MS, Groenwold RH, Berkelbach van der Sprenkel JW, *et al*. Automated detection of external ventricular and lumbar drain-related meningitis using laboratory and microbiology results and medication data. *PLoS ONE* 2011;6: e22846.

39 Pokorny L, Rovira A, Martin-Baranera M, *et al*. Automatic detection of patients with nosocomial infection by a computer-based surveillance system: a validation study in a general hospital. *Infect Control Hosp Epidemiol* 2006;27:500–3.

40 Chang YJ, Yeh ML, Li YC, *et al*. Predicting hospital-acquired infections by scoring system with simple parameters. *PLoS ONE* 2011;6:e23137.

41 Gastmeier P, Kampf G, Hauer T, *et al*. Experience with two validation methods in a prevalence survey on nosocomial infections. *Infect Control Hosp Epidemiol* 1998;19:668–73.

42 Jha AK, DesRoches CM, Kralovec PD, *et al*. A progress report on electronic health records in U.S. hospitals. *Health Aff (Millwood)* 2010;29:1951–7.

43 Giacometti A, Cirioni O, Schimizzi AM, *et al*. Epidemiology and microbiology of surgical wound infections. *J Clin Microbiol* 2000;38:918–22.

44 Horan TC, White JW, Jarvis WR, *et al*. Nosocomial infection surveillance, 1984. *MMWR CDC Surveill Summ* 1986;35:17SS–29SS.

45 Gastmeier P, Brauer H, Hauer T, *et al*. How many nosocomial infections are missed if identification is restricted to patients with either microbiology reports or antibiotic administration? *Infect Control Hosp Epidemiol* 1999;20:124–7.

46 Wright SB, Huskins WC, Dokholyan RS, *et al*. Administrative databases provide inaccurate data for surveillance of long-term central venous catheter-associated infections. *Infect Control Hosp Epidemiol* 2003;24:946–9.

47 Stevenson KB, Khan Y, Dickman J, *et al*. Administrative coding data, compared with CDC/NHSN criteria, are poor indicators of health care-associated infections. *Am J Infect Control* 2008;36:155–64.

48 Platt R, Yokoe DS, Sands KE. Automated methods for surveillance of surgical site infections. *Emerg Infect Dis* 2001;7:212–6.

49 Horan TC, Andrus M, Dudeck MA. CDC/NHSN surveillance definition of health care-associated infection and criteria for specific types of infections in the acute care setting. *Am J Infect Control* 2008;36:309–32.

50 European Centre for Disease Prevention and Control. *HAIICU protocol v1.01 standard and light*. Stockholm, Sweden; 2010. http://www.ecdc.europa.eu/en/aboutus/calls/Procurement%20Related%20Documents/5_ECDC_HAIICU_protocol_v1_1.pdf (accessed 10/06/2013).

51 Tinoco A, Evans RS, Staes CJ, *et al*. Comparison of computerized surveillance and manual chart review for adverse events. *J Am Med Inform Assoc* 2011;18:491–7.

52 Leal J, Laupland KB. Validity of electronic surveillance systems: a systematic review. *J Hosp Infect* 2008;69:220–9.

53 Freeman R, Moore LS, Garcia Alvarez L, *et al*. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. *J Hosp Infect* 2013;84:106–19.