Data elements and validation methods used for electronic surveillance of health care-associated infections: A systematic review

Kenrick D. Cato PhD, RN a, *, Bevin Cohen MPH a, b, Elaine Larson PhD, FAAN, RN, CIC a, b

a School of Nursing, Columbia University, New York, NY
b Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY

Key Words:
Automated surveillance
Automation/methods
Electronic Health Records

Background: We describe the primary data sources, data elements, and validation methods currently used in electronic surveillance systems (ESS) for identification and surveillance of health care-associated infections (HAIs), and compares these data elements and validation methods with recommended standards.

Methods: Using Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, a PubMed and manual search was conducted to identify research articles describing ESS for identification and surveillance of HAIs published January 1, 2009-August 31, 2014. Selected articles were evaluated to determine what data elements and validation methods were included.

Results: Among the 509 articles identified in the original literature search, 30 met the inclusion criteria. Whereas the majority of studies (83%) used recommended data sources and validated the numerator (80%), only 10% of studies performed external and internal validation. In addition, there was variation in the ESS data formats used.

Conclusions: Our findings suggest that the majority of ESS for HAI surveillance use standard definitions, but the lack of widespread internal data, denominator, and external validation in these systems reduces the reliability of their findings. Additionally, advanced programming skills are required to create, implement, and maintain these systems and to reduce the variability in data formats.

Copyright © 2015 by the Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.
used the electronic health record as solely a means for conducting chart review were excluded, as were those that investigated predictive risk modeling for HAI. We used the filters human, abstract, and English language on all searches. Table 1 summarizes the PubMed search query.

### Assessment of studies

To ensure articles matched the eligibility criteria, titles and abstracts were evaluated independently by each author and discrepant cases were settled by consensus. Full texts of the remaining articles were then reviewed independently by each author to verify that they met the inclusion criteria. After articles that met the inclusion criteria were identified, they were assessed using a modified framework originally developed by Woeltje et al.⁸ The first 3 articles were independently assessed and then discussed by all 3 authors, and any discrepancies were resolved by consensus. The remaining articles were abstracted by 1 of the 3 authors using an assessment framework.

The assessment framework developed by Woeltje et al.⁸ has 2 main components. First, for each of 5 types of infection, recommended data elements for ESS were outlined based on National Healthcare Safety Network definitions for HAI surveillance (http://www.cdc.gov/nhsn/pdfs/pscmanual/17pscnosinfdef_current.pdf). These included central line–associated bloodstream infection, catheter-associated urinary tract infection, surgical site infection, ventilator-associated pneumonia, and pneumonia to the list of HAIs because these were also investigated in the articles we reviewed. Second, 4 key concepts for describing data validation were recommended: internal and external validation and validation of numerator and denominator.⁸ Based on this framework, we evaluated each article to determine whether all recommended data elements were included and whether recommended validations were performed. The Woeltje et al.⁸ framework was modified only for surgical site infections, for which it was decided that an ESS would not require both procedure and diagnostic codes because there is considerable overlap between

---

**Table 1**

Search term used for the PubMed query

<table>
<thead>
<tr>
<th>Search term</th>
</tr>
</thead>
</table>

---

**Fig 1.** Flow chart used to identify articles examining automated healthcare-associated infection surveillance systems to ultimately select articles for inclusion in our analysis.
<table>
<thead>
<tr>
<th>Group</th>
<th>Article</th>
<th>Study setting and size</th>
<th>Health care-associated infection type</th>
<th>Data sources used</th>
<th>Used recommended data sources</th>
<th>Validation performed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benoit et al 2011</td>
<td>34 Hospitals, 4,585 patients, 12 states</td>
<td>Clostridium difficile</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Branch-Ellman et al 2014</td>
<td>5 Acute care VA hospitals</td>
<td>MDRO</td>
<td>Microbiology cultures</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Choudhari et al 2014</td>
<td>Teaching hospital 136 patients</td>
<td>CAUTI</td>
<td>Microbiology cultures, ADT, vital signs, urinalysis</td>
<td>Yes</td>
<td>Internal, numerator, denominator</td>
</tr>
<tr>
<td></td>
<td>De Bus et al 2014</td>
<td>University hospital 14-bed MICU, 22-bed SCU, 876 admissions</td>
<td>UTI, BSI</td>
<td>Microbiology cultures, ADT, vital signs, urinalysis</td>
<td>Yes</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>FitzHenry et al 2013</td>
<td>4 Hospitals</td>
<td>Clostridium difficile</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Harron et al 2013</td>
<td>109,654 Pediatric inpatient records</td>
<td>UTI</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>Inacio et al. 2011</td>
<td>Academic medical center, 204-bed ICU</td>
<td>UTI</td>
<td>Microbiology cultures, ADT, hospital procedure codes, hospital diagnosis codes</td>
<td>Yes</td>
<td>Numerator, denominator</td>
</tr>
<tr>
<td></td>
<td>Klein et al 2014</td>
<td>Tertiary referral centers, 2,080 patients</td>
<td>VAP, VAE</td>
<td>Microbiology cultures, ADT, vital signs, presence of endotracheal intubation device, laboratory (white blood cell count), ventilator settings (PEEP, FiO₂ antimicrobial use)</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Leal et al 2010</td>
<td>Academic safety net hospital, 2,449 surgical procedures</td>
<td>SSI</td>
<td>Microbiology cultures, ADT, antibiotic administration, hospital procedure codes, hospital diagnosis codes</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Leth et al 2010</td>
<td>Regional hospital, 306 patients</td>
<td>UTI</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Lo et al 2013</td>
<td>Tertiary care teaching hospital, 730-bed</td>
<td>CAUTI</td>
<td>Microbiology cultures, ADT, vital signs, urinalysis</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Peterson et al 2012</td>
<td>Health system, more than 300,000 patients</td>
<td>MDRO</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Schmiedeskamp et al 2009</td>
<td>23,520 Inpatient discharges</td>
<td>Clostridium difficile</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Shepard et al 2014</td>
<td>A 583-bed adult tertiary care center, 6,379 urine cultures</td>
<td>CAUTI</td>
<td>Microbiology cultures, ADT, vital signs, urinary catheter present, urinalysis</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Venable et al 2009</td>
<td>A 420-bed teaching hospital</td>
<td>CLABSI, CAUTI</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Wald et al 2014</td>
<td>A 425-bed university hospital, 1,695 patient visits</td>
<td>CAUTI</td>
<td>Microbiology cultures, ADT, urinary catheter present, urinalysis</td>
<td>Yes</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>Woeltje et al 2011</td>
<td>A 1,250-bed tertiary care teaching hospital</td>
<td>CLABSI</td>
<td>Microbiology cultures, ADT</td>
<td>No</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>Tseng et al 2012</td>
<td>A 2,200-bed teaching hospital</td>
<td>MDRO</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Numerator</td>
</tr>
<tr>
<td></td>
<td>Tseng et al 2013</td>
<td>A 2,200-bed teaching hospital</td>
<td>BS1</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>Wu et al 2009</td>
<td>A 2,200-bed teaching hospital</td>
<td>MDRO</td>
<td>Microbiology cultures, ADT</td>
<td>Yes</td>
<td>Internal, numerator</td>
</tr>
<tr>
<td></td>
<td>Apte et al 2011</td>
<td>Academic medical center, 28,956 hospital admissions</td>
<td>SSI</td>
<td>Microbiology cultures, ADT, antibiotic administration, hospital procedure codes</td>
<td>No</td>
<td>Internal, numerator, denominator</td>
</tr>
<tr>
<td></td>
<td>Apte et al 2011</td>
<td>Academic medical center, 320,000 inpatient discharges</td>
<td>BS1, PNU, SSI, UTI</td>
<td>Microbiology cultures, ADT, urinary catheter present, urinalysis, central venous catheter presence, antibiotic administration, hospital procedure codes</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Landers et al 2010</td>
<td>Academic medical center, 33,834 hospital admissions</td>
<td>UTI</td>
<td>Microbiology cultures, ADT, urinalysis</td>
<td>No</td>
<td>None</td>
</tr>
</tbody>
</table>
the 2; thus, this data element was considered present if either type of code was used.

RESULTS

As Figure 1 illustrates, 509 articles were initially identified (Table 1 lists the full search text). After removing duplicate citations and limiting articles to those with available abstracts, 383 abstracts were screened. An additional 77 were excluded during title and abstract review, primarily because they did not pertain to automated ESS. Full-text assessment of 35 articles resulted in 30 final studies that met inclusion criteria.

Table 2 provides a summary of each study reviewed, which included an array of HAIs: bloodstream infection = 10 and central line-associated bloodstream infection = 5, urinary tract infection = 7 and catheter-associated urinary tract infection = 7, surgical site infection = 5, multidrug-resistant organisms = 3, any ventilator-associated events = 1 and pneumonia = 2, and C difficile = 3. The majority of studies, 83%, used the recommended HAI-specific data sources in their ESS.

The articles reviewed did not always report how clinical facts (eg, laboratory results, diagnosis, and medications) were annotated and the corresponding vocabularies used to format the related data. However, there was variation in data formats for the studies that did provide a detailed description of data used by their ESS. These formats varied from unstructured, noncoded, and institution-specific coded data to internationally and nationally adopted formats like ICD-9. To determine antibiotics administered, textual medication names and institution-specific code forms were used. ICD-9,10,11,13,14,17 Systematized Nomenclature of Medicine—Clinical Terms,15-22 and free text from notes were used to determine hospital billing diagnosis and procedures.10,11,17,21 Microbiology results were formatted in institution-specific codes,10,11,13,14,17,22 and Logical Observation Identifiers Names and Codes.22 Validation performed

Validation of the numerator was performed most often (80%; 24 out of 30 studies). Checking of the actual data with an independent external organization validate the ESS, and the lack of validation of the surveillance systems. The increased number of data sources used in ESS reflects the fact that more electronic clinical data continue to become available.6 However, the fact that 17% of the studies in our review did not use the recommended data suggests that availability of relevant data is still 1 gap that must be filled before fully automated surveillance of HAI is possible.

The fact that a majority of studies used the recommended data elements to define an HAI is encouraging. Still, it is important to note that, for example, although all of the ESS used microbiology laboratory results, there was great variability in the structure of the
actual results. This lack of standardization of data is an impediment to having ESS that can be implemented uniformly across settings. Furthermore, lack of uniformity in data input increases the complexity of the systems and the required resources to implement and maintain these systems.

In our review most of the studies were conducted in academic medical centers, Veterans Administration hospitals, and 1 large health maintenance organization. This finding reflects the fact that only institutions with considerable financial resources can afford to implement these systems. Moreover, ESS are complex and range in sophistication, but in general they require high-level programming and technologic support. Clearly, the specialized workforce that is required to support these systems is sorely lacking. The creation, implementation, and management of ESS systems require individuals who understand the nuances of clinical data and analytical techniques and have the ability to extract and transform standard and nonstandard data sources.

None of the studies in our review met the criteria of having internal, external, denominator, and numerator validation. It is important to note that denominator validation is often not applicable and feasible to calculate for the HAIIs that are applicable to an entire inpatient population. Whereas the logistics of accomplishing the task of validating the data are daunting, validation is vital and must be performed on these systems. Research has indicated that ESS studies that have performed the requisite numerator, denominator, and/or external validation found high variability in sensitivity/specificity. In addition, studies have also highlighted the lack of completeness and bias of electronic patient data.

Our literature review was limited by the inclusion of only articles written in English with abstracts, and the use of a single database with a limited number of search terms. Therefore, some articles pertaining to the topic could have been missed. Nonetheless, our findings present the state of the science in ESS research and point to important future directions for obtaining investigation. We recommend that future ESS HAI surveillance focus on obtaining high-quality data, employing dedicated programmers with advanced skills, and performing more thorough validation.

References
