Mathematica
Progress Together

Memo

To: Alyson Schuster, Dianne Feeney, Andi Zumbrum, and Zach Goldberg
From: Harsha Amaravadi, Anthony Austin, and Bailey Orshan
Date: 4/13/2021
Subject: HAI Measure Environmental Scan

HAI Measure Environmental Scan

Purpose

The Maryland Health Services Cost Review Commission (HSCRC) requested an environmental scan that assesses and summarizes current literature on surveillance metrics related to the effectiveness of the following HAI measures:

- Catheter-Associated Urinary Tract Infection (CAUTI),
- Central Line-Associated Blood Stream Infection (CLABSI),
- Clostridium difficile Infection, *C. difficile* (CDI),
- Methicillin-Resistant Staphylococcus aureus Bacteremia (MRSA), and
- Surgical Site Infection (SSI) for colon procedures and abdominal hysterectomies

Parameters such as validity, reliability, sensitivity, specificity, and reporting rates were examined to contextualize the use of the HAI measures when determining quality.

Methods

Mathematica performed this literature review in multiple steps. First, by performing targeted searches of peer-reviewed journals, we identified a set of articles related to the HAI measures that included both systematic reviews and primary analyses. We limited the search to literature that included primary analyses on disease surveillance metrics (e.g., sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and reviews that addressed issues related to the reporting of measures. We then identified key themes across the articles and measures, summarizing the overarching findings regarding surveillance metrics and inconsistency in measurement. Below is a table defining the key surveillance metrics examined in this environmental scan.

Table 1: Definition of surveillance metrics

Surveillance Metric	Definition
Sensitivity	Probability of testing positive when infection present
Specificity	Probability of testing negative when infection absent
Positive Predictive Value (PPV)	Probability of a patient having infection when the test result is positive
Negative Predictive Value (NPV)	Probability of a patient not having infection when the test result is negative

Summary of Findings

CAUTI and CLABSI surveillance validation studies are the most abundant in NHSN surveillance literature. The literature reveals that misclassification of these conditions is variable.

We identified studies, both systematic reviews and primary analyses that assessed sensitivity, specificity, PPV, and NPV of **CAUTI and CLABSI and observe that both measures are generally underreported**. See Table 2 for a summary of the findings.

Table 2: Results of Validity Test across HAI measures (sensitivity, specificity, PPV, NPV)

Author, year	Measure	Study Design	Data Collection Period	Results
Larsen, 2019	CLABSI	Review of cohort studies with publicly reported CLABSI rates N=7,160 participants in 9 retrospective cohort studies	Cohort studies from January 2008 - 2017	Specificity ranged from: **0.70** CI [0.58;0.81] **0.99** CI [0.99;1.00] Sensitivity ranged from: **0.42** [0.15;0.72] **0.88** [0.77;0.95]
Bagchi, 2018	CLABSI	Individual data review of state health department records N=23 state health departments,	Validations conducted from 2008-2016	Pooled specificity: **0.985** [0.931;1.00] Pooled sensitivity: **0.845** [0.324;1.00] Pooled PPV: **0.941** [0.87;1.00] Pooled NPV: **0.959** [0.924;0.99]
Thompson, 2013	CLABSI	Retrospective medical record review in 6 New Mexico hospitals N=123 participants	Hospitalizations from 2009-2010	Specificity: **1.00** Sensitivity: **0.667** PPV: **1.00** NPV: **0.965**

Bagchi, 2019	CAUTI	Individual data review of state health department records N=50 state health departments, 7,073 participants	Validations conducted from 2010-2017	Pooled specificity: **0.988** [0.90;1.00] Pooled sensitivity: **0.883** [0.50;1.00] Pooled PPV: **0.936** [0.763;1.00] Pooled NPV: **0.976** [0.921;0.998]
Hanna, 2013	CAUTI	Medical record review assessing i) clinician diagnosis of CAUTI compared to NHSN definition (N=387) ii) clinician diagnosis of CAUTI compared to infectious disease consultation (N=211)	2010-2011	*(clinician diagnosis compared to NHSN)* Specificity: **0.662** Sensitivity: **0.470** PPV: **0.343** NPV: **0.737** *(clinician diagnosis compared infectious disease consult* Specificity: **1.00** Sensitivity: **0.574** PPV: **0.50** NPV: **1.00**

Overall, the findings reveal that CLABSI sensitivity rates are consistently rated as "moderate", while specificity is generally high across studies. There is more consistency in CAUTI sensitivity and specificity rates, and both are generally high (Larsen, 2019; Bagchi, 2018). Table 3 summarizes the reasons identified in the literature scan for CAUTI and CLABSI misclassification and inconsistent reporting.

The Larsen, 2019 systematic review examined nine cohort studies that compared publicly reported CLABSI rates with those of an expertly trained reviewer. The study objective was to establish reliability of CLABSI surveillance metrics. The reviewers found that publicly reported CLABSI rates were more likely to be underreported than overreported. In addition, in all nine studies, CLABSI sensitivity was lower than specificity, suggesting that a misclassification of CLABSI was more likely to be a false negative than a false positive.

Bagchi et al. conducted two similar primary analyses on HAI rates in state health departments in the United States, examining CLABSI in 2018 and CAUTI in 2019. The study design involved aggregating results from state health department validation studies, which looked at measure sensitivity, specificity, PPV, and NPV. The Bagchi 2019 CAUTI study also examined changes in surveillance metrics before and after the NHSN changed the CAUTI definition framework in 2015. They observed that the error rate declined significantly from 4.3% to 2.4% after the change in surveillance metrics. Changing definitions can increase accuracy, but can also cause confusion in the clinical community, and should be considered when interpreting quality scores across facilities. For example, Bagchi et al. note that misapplication of a general NHSN definition is the primary reason for CAUTI misclassification, and cite examples where facilities used the date of

urine culture (pre-2015 definition) as opposed to the date of the first element as the date of event (post-2015 definition).

Thompson et al. conducted a primary analysis, via a retrospective medical record review, that assessed CLABSI surveillance metrics by comparing clinician assessment of infection with NHSN definitions. The study consisted of 123 patient cases across six New Mexico hospitals and found that specificity, PPV, and NPV were quite high, yet sensitivity was moderate, suggesting that CLABSI was underreported.

Table 3: Reasons for CAUTI and CLABSI Misclassification

Author, year and Description	Reasons for infection misclassification or inconsistent reporting
Bagchi, 2018 and Bagchi, 2019 Retrospective cohort studies on CLABSI and CAUTI misclassification in state health departments **Larsen, 2019** Review of cohort studies with publicly reported CLABSI rates	• misapplication of NHSN CAUTI/CLABSI definition • missed case findings • misapplication of general NHSN HAI definition • application of clinical judgment over surveillance definition, including subjective clinician reporting • inadequate physician education • insufficient hospital resources

Abundance of HAI Surveillance literature varies by measure

CAUTI and CLABSI surveillance literature is more common than for other HAIs, however, in 2020, Bagchi et al. continued their state health department analyses on the full suite of HAI measures and presented preliminary findings in an oral presentation. (Bagchi, 2020). This presentation provided a few more surveillance metric results on MRSA, CDI, and SSI, for which there are limited primary analyses on sensitivity, specificity, PPV, and NPV in the literature. They found that:

- the pooled mean sensitivity for colon SSI, was lowest across all HAI measures, at 0.731,
- the pooled mean sensitivity for CDI was the highest across all HAI measures, at 0.927,
- measures utilizing "Laboratory Identified" (LabID) surveillance systems (MRSA and CDI), had the lowest NPV across all HAI measures (0.588 and 0.551, respectively)
- MRSA also had the highest error rate at 13.6%, and
- Abdominal hysterectomy SSI had the lowest error rate at 2.5%.

They also cited the primary reasons for missed MRSA and CDI case findings as failure to review candidate events and gaps in understanding the 14-day reporting rule of LabID protocol.

CDI measurement may be overreported at the hospital level

In 2013, the NHSN established a new surveillance method, called "laboratory-identified" (LabID) CDI events, which uses electronically captured laboratory data and hospital admission dates to determine hospital-onset (HO) versus community-onset (CO) surveillance categories (Durkin, 2015). Albert et al. conducted a retrospective medical chart review study examining 212 CDI cases within an urban medical center that were all initially considered to be hospital acquired based on LabID event definition. Retrospective chart review, which incorporated clinical data in addition to the existing molecular testing, was utilized to more accurately re-classify the cases as either: CO- CDI, recurrent/relapse CDI, asymptomatic colonization, colonization with limited symptoms, possible HO-CDI, and probable HO-CDI. Results found that only 38.7% of the 212 cases were possible HO-CDI, and 24.5% were probable HO-CDI. Accordingly, the researchers suggest that new CDI LabID surveillance definitions may actually lead to overreporting of HO-infection, as about 37% of the suspected hospital-acquired cases were not considered possible or probable HO-CDI post review. They explain that though LabID methods simplify CDI surveillance, the lack of clinical assessment can lead to misclassification. They note that about 10-15% of patients may enter a hospital with asymptomatic *C.difficile,* which is difficult to distinguish from HO-CDI from molecular testing alone. Based on the 2013 results, Albert et al. cite laxative use and failure to identify community-onset infection as sources of misclassified HO-CDI. (Albert, 2018). Durkin et al. also came to a similar conclusion when they conducted a cohort study in 29 hospitals in 2013, which compared LabID identified cases against cases identified through traditional surveillance methods. In aggregate, they found significant differences in HO-CDI rates based on the surveillance method. They observed that LabID identified HO-CDI rate was 6.0 per 10,000 patient-days, vs 4.4 per 10,000 patient-days based on traditional surveillance, again suggesting overreporting of HO-CDI.

Several studies indicate that surveillance definitions and clinical practice definitions differ, suggesting that further clinician education and auditing interventions need to be consistently applied for fair comparisons.

There have been several accounts of inconsistent application of surveillance criteria within the infection prevention clinical community (Wright, 2017). According to Talbot et al. clinical diagnoses will inherently have some level of subjectivity and are used to guide treatment of individual patients; however surveillance definitions should be more standardized with the intent to assess HAI burden across healthcare settings and to quantify the effect of prevention efforts (Talbot, 2013). Van Mourik et al. make the nuanced point, however, that there is always a tradeoff in developing surveillance systems: increased standardization afforded through methods like automation sometimes come at the cost of decreased clinical relevance. They suggest that

surveillance methods be chosen based on the downstream goals of the surveillance (e.g., pay-for-performance, general hospital-based quality improvement, or research) (Van Mourik, 2018).

Several HAI conditions also have measurement systems and scales outside of the NHSN that have varying degrees of difference with NHSN definitions. For example, SSI rates are also monitored using the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP). A study conducted by Ju et al. examined differences in SSI colon rates at 16 hospitals when using NHSN and ACS NSQIP criteria. They found significant differences between the methods, noting that ACS NSQIP rates were always higher than NHSN rates, and that SSI colon cases that were managed outpatient were more likely to be missed under NHSN criteria. (Ju, 2015). Thompson et al. also underscore this sentiment, noting that further alignment between clinical and surveillance definitions will also garner additional "buy-in" from medical staff, which can in turn influence more accurate application of the criteria (Thompson, 2013).

Wright et al. conclude that differences in clinical and surveillance definitions among physicians and infection preventionists yield differences in classification of disease according to NHSN definitions. They conducted a study from 2010 – 2016, in which a variety of clinicians (including physicians and infection preventionists) answered survey questions to test their application of standardized NHSN criteria to case study reports to classify disease. In aggregate, study participants (n=7,950) made correct decisions (according to NHSN gold standard) only 62.5% of the time. The study underscores the importance of ensuring that correct NHSN criteria are applied in practice in all healthcare settings. Further education and auditing initiatives can attempt to rectify such differences (Wright, 2017 & Hamazay, 2013).

Similar findings were observed in a Michigan-based hospital study conducted by Hanna et al. The study compared clinician diagnoses, infectious disease consultation results, and NHSN CAUTI criteria definitions as applied to possible CAUTI incidents between 2010 and 2011 (n=387 patients). Results showed that clinical practice definitions and surveillance definitions can vary greatly, further emphasizing the importance of interpreting NHSN surveillance data when assessing facility quality. The study also reported significant variability in sensitivity, specificity, PPV, and NPV when comparing each definition (physician diagnosis, infectious disease consultation, or NHSN definition) as the gold standard. See Table 2 for sensitivity and specificity results. The authors also commented on ways to interpret NHSN CAUTI rates, noting that the rates may not be reflective of patient outcomes related to complications, but should be looked as markers of improved infection prevention (Hanna, 2013).

Finally, Hazamy et al. conducted a study among New York State hospitals to investigate the effect of hospital auditing on surveillance metrics for CLABSI. They found that between 2007 and 2010, specificity increased from 90 to 99%, while sensitivity remained relatively stable at 71%, and overall CLABSI rate increased by 5.6%. Their research suggests that auditing programs increase reporting and accuracy of surveillance measures; however, auditing programs are not consistently implemented across the country or across various hospitals. Auditing and

education programs can bridge some gaps between clinical and NHSN definition application and can increase reporting rates that in turn affect surveillance metric scores.

HAI measures are susceptible to surveillance bias, which should be considered when assessing quality across facilities.

Surveillance bias is a non-random form of information bias that leads to increased case findings simply because a facility is looking for cases (Haut, 2011). Surveillance bias may be the source of differences in outcomes across facilities rather than true differences in quality of care. In a 2010 study conducted by Yin et al., computer automated surveillance of CLABSI was compared against traditional surveillance methods conducted by infection preventionists. Across the 20 intensive care units assessed, the median CLABSI rate increased about three-fold when using computer automated surveillance. The results suggest that surveillance bias could be source of the observed variation in CLABSI surveillance, which "may complicate interinstitutional comparisons of publicly reported central line–associated BSI rates" (Yin, 2010). Similarly, Niedner et al. conducted a survey-based study in pediatric intensive care units in which a "surveillance aggressiveness" score was calculated to quantify the robustness of a facility's surveillance practices. The score was calculated by counting events believed to favor identification of CLABSI, essentially making the score a marker of surveillance bias. There was a statistically significant correlation between the surveillance aggressiveness score and the CLABSI rate observed at each facility, suggesting surveillance bias may have occurred. (Niedner, 2010). For this reason, Van Mourik et al. note that declines in HAI rates cannot always be taken at face value as objective increases in quality, without first considering surveillance bias and possible underreporting (Van Mourik, 2018).

Conclusion

The literature surveyed includes both reviews and primary analysis of surveillance metrics for HAI measures across a variety of healthcare entities. Though study designs and abundance of literature on the different HAI measures vary, there is universal commentary that HAI rates vary across facilities in part because of differences in the application of NHSN criteria,clinical definitions, and surveillance bias. Auditing and clinical education can help bring alignment to these differences to reduce overreporting and underreporting of HAIs.

References

Albert, Kathryn, et al. "Overreporting healthcare-associated C. difficile: A comparison of NHSN LabID with clinical surveillance definitions in the era of molecular testing." *American journal of infection control* 46.9 (2018): 998-1002.

Bagchi, Suparna, Bonnie Norrick, and Nigel Lewis. "Identifying Opportunities to Improve Accuracy of NHSN Reporting: Lessons Learned From State Health Department Validations." *Infection Control & Hospital Epidemiology* 41.S1 (2020): s41-s41.

Bagchi, Suparna, et al. "Accuracy of catheter-associated urinary tract infections reported to the National Healthcare Safety Network, January 2010 through July 2018." *American journal of infection control* 48.2 (2020): 207-211.

Bagchi, Suparna, et al. "State health department validations of central line–associated bloodstream infection events reported via the National Healthcare Safety Network." *American journal of infection control* 46.11 (2018): 1290-1295.

Durkin, Michael J., et al. "A comparison between National Healthcare Safety Network laboratory-identified event reporting versus traditional surveillance for Clostridium difficile infection." *Infection control and hospital epidemiology* 36.2 (2015): 125.

Hanna, Fadi Al-Qas, et al. "Clinician practice and the National Healthcare Safety Network definition for the diagnosis of catheter-associated urinary tract infection." *American journal of infection control* 41.12 (2013): 1173-1177.

Hazamy, Peggy Ann, et al. "Trends in validity of central line–associated bloodstream infection surveillance data, New York State, 2007-2010." *American journal of infection control* 41.12 (2013): 1200-1204.

Haut, Elliott R., and Peter J. Pronovost. "Surveillance bias in outcomes reporting." *Jama* 305.23 (2011): 2462-2463.

Ju, Mila H., et al. "A comparison of 2 surgical site infection monitoring systems." *JAMA surgery* 150.1 (2015): 51-57.

Larsen, Emily N., et al. "A systematic review of central-line–associated bloodstream infection (CLABSI) diagnostic reliability and error." *Infection Control & Hospital Epidemiology* 40.10 (2019): 1100-1106.

Lin MY, Hota B, Khan YM, et al. Quality of Traditional Surveillance for Public Reporting of Nosocomial Bloodstream Infection Rates. *JAMA.* 2010;304(18):2035–2041. doi:10.1001/jama.2010.1637

Niedner, Matthew F. "The harder you look, the more you find: catheter-associated bloodstream infection surveillance variability." *American journal of infection control* 38.8 (2010): 585-595.

Talbot, Thomas R., et al. "Public reporting of health care–associated surveillance data: recommendations from the Healthcare Infection Control Practices Advisory Committee." *Annals of internal medicine* 159.9 (2013): 631-635.

Thompson, Deborah L., Monear Makvandi, and Joan Baumbach. "Validation of central line–associated bloodstream infection data in a voluntary reporting state: New Mexico." *American journal of infection control* 41.2 (2013): 122-125.

Van Mourik, Maaike SM, et al. "Designing surveillance of healthcare-associated infections in the era of automation and reporting mandates." *Clinical Infectious Diseases* 66.6 (2018): 970-976.

Wright, Marc-Oliver, Katherine Allen-Bridson, and Joan N. Hebden. "Assessment of the accuracy and consistency in the application of standardized surveillance definitions: A summary of the American Journal of Infection Control and National Healthcare Safety Network case studies, 2010-2016." *American journal of infection control* 45.6 (2017): 607-611.