COMMENTARY

Building Data Quality and Confidence in Data Reported to the National Healthcare Safety Network

Kathryn E. Arnold, MD;[1] Nicola D. Thompson, PhD, MS[1]

(See the article by Oh et al, on pages 439–445.)

Since 2009, the Centers for Disease Control and Prevention (CDC) has provided funding to 51 states and territories to develop healthcare-associated infection (HAI) prevention programs and, in some states, funds for systematic validation of data reported by hospitals to the National Healthcare Safety Network (NHSN). As of 2011, at least 11 state health departments (SHDs) have completed validation projects for central line–associated bloodstream infections (CLABSIs), and in several states validation is ongoing. Differences in resource availability and the numbers of hospitals reporting to NHSN within states have resulted in a variety of methods used to validate HAI data. To enable efficient use of resources, SHDs often employ a targeted approach to selecting the healthcare facilities, units, and medical charts for validation;[1] therefore, direct comparison of the results may be limited. Nevertheless, common themes include an opportunity for hospital staff to receive additional education and training on NHSN methods, and impartial feedback regarding the quality and completeness of their HAI reporting efforts. Both activities build confidence in the data available to SHDs, group users, and the CDC to inform policy and HAI prevention efforts.

The study by Oregon investigators in this issue[2] represents the second peer-reviewed publication reporting results of SHD validation of CLABSI data (results from several other efforts have been reported at scientific meetings or on SHD Web sites). Using comparable methodology to identify hospitals, units, and medical charts for validation, the results of the two published studies[2,3] report high specificity (99%) and positive-predictive (85%–92%) and negative-predictive (94%–98%) values for CLABSI reporting to NHSN by acute care hospitals. Both, however, also identified CLABSI underreporting. Unlike Connecticut, which reported an estimated sensitivity of 48% for CLABSI reporting,[3] the Oregon validation study reported a much higher sensitivity of 72% (95% confidence interval, 62%–81%).[2] Moreover, investigators reported 100% sensitivity at most (75%) of the 44 hospitals participating in the validation effort. The results of both studies highlight the importance of performing CLABSI validation to assess the quality of data reported to NHSN and to build confidence in facility-specific and statewide measurement of CLABSIs. Differences between the results of these two assessments may relate to actual differences in the accuracy of reporting and surveillance bias,[4] but they are as likely to be due to differences in validation methodology.

An important distinction between the validation methodologies used by the two SHDs was the process used to assign the validated CLABSI status, used as the referent standard for calculating sensitivity and specificity of hospital CLABSI reporting. In Connecticut, validated CLABSI status was assigned solely by SHD investigators.[2] By comparison, Oregon investigators engaged hospital staff in a review process when the CLABSI status assigned by the hospital was discordant with that assigned by the SHD, enabling Oregon investigators to assure that all information relevant to applying the NHSN CLABSI definition was available to the SHD for determination of the validated CLABSI event. As suggested by Oh and colleagues,[2] this review process likely helped to eliminate data quality bias[5] that could otherwise result from challenges encountered when accessing and identifying information from heterogeneous electronic medical record systems. SHDs are uniquely positioned to engage facilities in collaborative validation reviews that allow for transparency, education, and relationship-building, which may not be possible in a regulatory audit process, and to achieve results that are ultimately more credible to all parties, including reporting hospitals, SHDs, and other users of NHSN data.

Although validation of CLABSI events (ie, numerators) has been a hallmark of SHD HAI validation programs, SHDs have also taken steps to validate CLABSI denominator data (patient-days and central line–days); the approach has generally been qualitative rather than quantitative.[3] In Oregon,

no CLABSI denominator data were validated, a limitation the authors readily acknowledge. Inaccurate CLABSI denominator data can result in either under- or overestimated rates, leading to inaccurate interfacility comparisons. Prior studies indicate a need to ensure that CLABSI denominator data reported to NHSN are valid, especially in facilities where electronic systems are utilized to obtain these data.[6,7] Because retrospective validation of CLABSI denominator data is challenging, innovative methods and improved efforts to ensure accurate denominator data, not only as part of SHD HAI validation programs but also prospectively by hospitals reporting data to NHSN, are necessary.

Several studies have identified interobserver variability in assigning CLABSI case status across infection preventionists and facilities.[8,9] Some components of the NHSN CLABSI case definition are vulnerable to subjective interpretation, such as distinguishing whether an infection may be community or healthcare associated, and legitimate differences of opinion may lead to challenges in reliable case classification. However, failure to adhere to objective NHSN surveillance criteria, such as those for classifying common commensals, and a preference for use of clinical rather than NHSN surveillance definitions have also been identified.[3,9,10] This variability can be reduced by increased delivery of NHSN user training and education, such as that provided by SHDs during HAI data validation initiatives. For example, the SHD in New York has reported annual improvements in agreement between hospital NHSN CLABSI reporting and SHD validation reviews between 2007 and 2009.[11]

New uses of NHSN data for public reporting and Centers for Medicare and Medicaid Services payment decisions create an imperative for assuring that hospitals and payers have confidence in the data reported to NHSN. The CDC is working to address this new imperative while trying to maintain and enhance the usefulness of NHSN data for its original purpose: real-time infection prevention and control. The CDC's efforts to assure and build confidence in NHSN CLABSI data quality are multifaceted:

- The CDC sponsors training to ensure that data are accurately and consistently collected and categorized. Formats include Web-based NHSN surveillance training modules (http://www.cdc.gov/nhsn/training.html), with webinars, slide sets, and new self-paced, interactive online training courses offering continuing-education credits upon successful completion of an assessment; provision of NHSN training during CDC-hosted events and at professional meetings and conferences; and SHD training and education activities conducted through prevention collaboratives and during HAI validation work.
- The CDC continues to refine the NHSN system to remove opportunities for human error. Examples include software changes such as business rules and cross-field edit checks to prevent data entry errors, system alerts to inform users of missing data, and the availability of data

quality reports to inform users of aberrant data. In addition, the CDC is exploring changes in methodology to minimize unreliable application of standard definitions and operations. Toward this end, the CDC is currently engaged with the Healthcare Infection Control Practices Advisory Committee and external partners to modify NHSN surveillance definitions and operations to increase clinical credibility and improve reliability in the era of public reporting.
- The CDC is developing standardized best practices for NHSN HAI data validation. Through exploring methods and tools developed by SHDs during their HAI validation work, the CDC has identified validation practices that can be integrated into the NHSN system as business rules or data quality reports and is developing guidance and tools for efficient validation work to be implemented by states as resources become available.

Several SHD HAI data validation initiatives, such as the one reported here, have clearly demonstrated their value in improving the quality of data reported to NHSN. While underreporting of CLABSIs by hospitals has been identified, CLABSI data reported by the majority of hospitals participating in these validation initiatives appear to be accurate, and ongoing SHD validation efforts have demonstrated that hospital CLABSI surveillance and reporting improve with the receipt of education and feedback over time. SHDs are uniquely positioned to provide oversight and perform validation of HAI data. Expansion of the scope and number of SHDs engaged in HAI data validation will help to build confidence in NHSN data that are now widely used for public reporting purposes.

REFERENCES

1. Kainer MA, Mitchell J, Frost BA, Soe MM. Validation of central line associated blood stream infection (CLABSI) data submitted to the National Healthcare Safety Network (NHSN): a pilot study by the Tennessee Department of Health (TDH). In: *Abstracts and scientific program, Fifth Decennial International Conference on Healthcare-Associated Infections.* Atlanta, GA, March 18–22, 2010. Abstract 456.

2. Oh JY, Cunningham MC, Beldavs ZG, et al. Statewide validation of hospital-reported central line–associated bloodstream infections: Oregon, 2009. *Infect Control Hosp Epidemiol* 2012;33(5): XXX–XXX (in this issue).

3. Backman LA, Melchreit R, Rodriguez R. Validation of the surveillance and reporting of central line-associated bloodstream infection data to a state health department. *Am J Infect Control* 2010;38(10):832–838.

4. Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA* 2011;305(23):2462–2463.

5. Terris DD, Litaker DG. Data quality bias: an underrecognized source of misclassification in pay-for-performance reporting? *Qual Manag Health Care* 2008;17(1):19–26.

6. Soe MM, Kainer MA. Sustainable, cost-effective internal data validation of healthcare associated infections surveillance reported to the National Healthcare Safety Network (NHSN). In: *Abstracts and scientific program, Fifth Decennial International Conference on Healthcare-Associated Infections.* Atlanta, GA, March 18–22, 2010. Abstract 81.

7. Chernetsky Tejador SG, Garrett G, Jacob J, et al. Electronic documentation of central line-days: validation is essential. In: *Abstracts and program of the 21st Annual Scientific Meeting of*

the Society for Healthcare Epidemiology of America (SHEA). Dallas, TX, April 1–4, 2011. Abstract 308.

8. Mayer J, Howell J, Green T, et al. Assessing inter-rater reliability (IRR) of surveillance decisions by infection preventionists (IPs). In: *Abstracts and scientific program, Fifth Decennial International Conference on Healthcare-Associated Infections.* Atlanta, GA, March 18–22, 2010. Abstract 79.

9. Malpiedi P, Hota B, Magill S, et al. Interobserver variability in bloodstream infection determinations using National Healthcare Safety Network definitions. In: *Abstracts and program of the 21st Annual Scientific Meeting of the Society for Healthcare Epidemiology of America (SHEA).* Dallas, TX, April 1–4, 2011. Abstract 305.

10. Mayer J, Weir CR, Greene T, et al. Use of electronic reporting and annotations by infection preventionists (IPs) to describe differences in the application of surveillance criteria. In: *Abstracts and program of the 21st Annual Scientific Meeting of the Society for Healthcare Epidemiology of America (SHEA).* Dallas, TX, April 1–4, 2011. Abstract 573.

11. Hazamy PA, VanAntwerpen C, Tserenpuntsag B, et al. Determining accuracy of CLABSI data submission for public reporting hospital-acquired infections in New York State from 2007–2009. *Am J Infect Control* 2011; 39(5):E132–E133.