ORIGINAL ARTICLE

Can National Healthcare-Associated Infections (HAIs) Data Differentiate Hospitals in the United States?

Max Masnick, PhD;[1] Daniel J. Morgan, MD, MS;[1,2] John D. Sorkin, MD, PhD;[3] Mark D. Macek, DDS, DrPH;[4] Jessica P. Brown, PhD;[1] Penny Rheingans, PhD;[5] Anthony D. Harris, MD, MPH[1] for the CDC Prevention Epicenter Program

OBJECTIVE. To determine whether patients using the Centers for Medicare and Medicaid Services (CMS) Hospital Compare website (http://medicare.gov/hospitalcompare) can use nationally reported healthcare-associated infection (HAI) data to differentiate hospitals.

DESIGN. Secondary analysis of publicly available HAI data for calendar year 2013.

METHODS. We assessed the availability of HAI data for geographically proximate hospitals (ie, hospitals within the same referral region) and then analyzed these data to determine whether they are useful to differentiate hospitals. We assessed data for the 6 HAIs reported by hospitals to the Centers for Disease Control and Prevention (CDC).

RESULTS. Data were analyzed for 4,561 hospitals representing 88% of registered community and federal government hospitals in the United States. Healthcare-associated infection data are only useful for comparing hospitals if they are available for multiple hospitals within a geographic region. We found that data availability differed by HAI. *Clostridium difficile* infections (CDI) data were most available, with 82% of geographic regions (ie, hospital referral regions) having >50% of hospitals reporting them. In contrast, 4% of geographic regions had >50% of member hospitals reporting surgical site infections (SSI) for hysterectomies, which had the lowest availability. The ability of HAI data to differentiate hospitals differed by HAI: 72% of hospital referral regions had at least 1 pair of hospitals with statistically different risk-adjusted CDI rates (SIRs), compared to 9% for SSI (hysterectomy).

CONCLUSIONS. HAI data generally are reported by enough hospitals to meet minimal criteria for useful comparisons in many geographic locations, though this varies by type of HAI. CDI and catheter-associated urinary tract infection (CAUTI) are more likely to differentiate hospitals than the other publicly reported HAIs.

Infect Control Hosp Epidemiol 2017;38:1167–1171

Healthcare-associated infection (HAI) data are reported by US hospitals to the Centers for Disease Control and Prevention (CDC) via the National Healthcare Safety Network (NHSN). These data are then reported to the public by the Centers for Medicare and Medicaid Services (CMS), with the goal of helping patients decide which hospitals to utilize. The CMS website, Hospital Compare (http://medicare.gov/hospitalcompare), allows the public to search for hospitals in a geographic region, and to then compare these hospitals based on a variety of hospital quality data including HAIs.

The CMS provides data for the following HAIs: (1) catheter-associated urinary tract infection (CAUTI); (2) central line-associated bloodstream infection (CLABSI); (3) surgical site infection (SSI), separately for colon surgery and hysterectomy; (4) methicillin-resistant *Staphylococcus aureus* (MRSA) bloodstream infection; and (5) *Clostridium difficile* infection (CDI).

The CMS uses standardized infection ratios (SIRs) to account for differences in underlying patient factors that increase HAI risk and can differ among hospitals. These SIRs are calculated by dividing the number of observed infections by the predicted number of infections for that hospital (see the Discussion section for details). The predicted numbers of infections are based on a limited set of risk characteristics that differ by HAI.[1]

Public reporting of HAI data requires substantial resources from hospitals[2] as well as additional resources from CDC and CMS to aggregate and publish these data. However, it is not known whether these data are useful in differentiating hospitals. For example, are there substantial differences in CAUTI rates among the hospitals in the Baltimore, Maryland, metropolitan area? If all hospitals in Baltimore have very low CAUTI rates, the CAUTI data are minimally informative when

Affiliations: 1. Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, Maryland; 2. Veterans Affairs Maryland Healthcare System, Baltimore, Maryland; 3. Veterans Affairs Maryland Healthcare System Geriatrics Research, Education, and Clinical Center, Baltimore, Maryland; 4. University of Maryland School of Dentistry, Baltimore, Maryland; 5. Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, Maryland.

Received February 20, 2017; accepted July 11, 2017

 0899-823X/2017/3810-0005. DOI: 10.1017/ice.2017.179

choosing among these hospitals. The aim of this study was to determine whether nationwide HAI data published by CMS are useful in differentiating hospitals within a geographic area.

METHODS

For our analysis, we used publicly available HAI data published by CMS in January 2015. These data cover calendar year 2013.[3] We analyzed these data from the perspective of someone searching for a hospital. To do this, we compared HAI data among geographically proximate hospitals that would constitute a plausible set of choices for an actual person looking for a hospital in their geographic region.

We defined geographic regions for grouping hospitals using hospital referral regions (HRRs) as specified by the Dartmouth Atlas.[4] The HRRs are geographic areas based on "where patients were referred for major cardiovascular surgical procedures and for neurosurgery."[4] Each hospital in the CMS data[3] was assigned to a single HRR based on the "providerid" field that appeared in both datasets. Hospital referral regions were identified for 97% of hospitals in the CMS dataset based on "providerid," and HRRs for the remainder were determined using zip codes.

Several criteria were assessed for each geographic region (HRR) as follows:

Outcome 1: Data availability. We determined the proportion of hospitals in each geographic region (HRR) where data for the HAI were available (ie, where data exist that can be used for comparing hospitals). If no HAI data are available for a high proportion of hospitals, then these data do not help differentiate hospitals.

Outcome 2: HAI performance diversity. We determined the percentage of hospitals in each HRR where at least 1 pair of hospitals had statistically different SIRs for the HAI, suggesting a difference in performance between the hospitals for that HAI. Then, SIR comparisons were performed using a method published by the CDC for comparing 2 SIRs, and a 2-tailed significance of $\sigma = 0.05$ was used.[5] If hospitals all perform similarly with in a HRR, then the HAI data are not useful in comparing hospitals in that geographic area. If this were true for many HRRs, then HAI data would not generally be useful for comparing hospitals nationwide.

RESULTS

Descriptive Results

Publicly available CMS HAI data were analyzed for all 4,561 hospitals in 306 HRRs (geographic regions). These data represent all hospitals available in the CMS data and 88% of 5,187 registered community and federal government hospitals in the United States at the time of the analysis.[6] The number of hospitals in each HRR ranged from 2 to 83, with a median of 11 (interquartile range, 6–17). Most HRRs included between 2 and 20 hospitals (Figure 1). Figure 2 shows the distribution of SIRs for each HAI within each HRR.

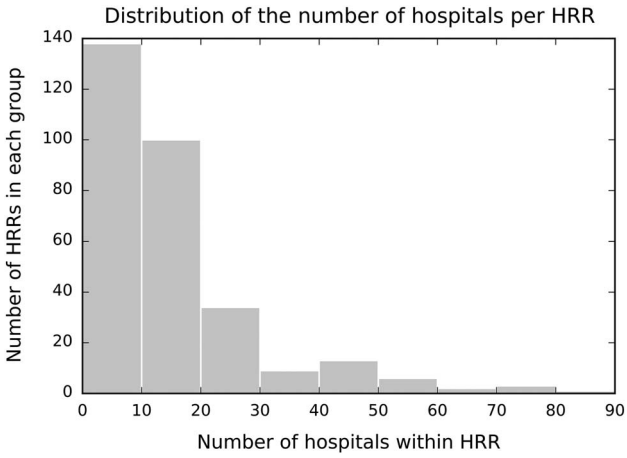

FIGURE 1. The distribution of the number of hospitals within each hospital referral region.

Outcome 1: Data availability. Table 1 shows the proportion of hospitals in each HRR with reported data, meaning that the hospital reports a number of infections rather than "not available" for a given HAI. Data that are "not available" indicate that either (1) the hospital did not report the data or (2) the denominator (eg, central line days for CLABSI) is below a CMS reporting threshold. (The threshold is not to our knowledge specified by CMS.[3]) For *C. difficile*, 250 of the 306 HRRs (82%) had at least 50% of member hospitals reporting data; 47% of HRRs had at least 50% of hospitals reporting CAUTI data; 43% of HRRs had at least 50% of hospitals reporting SSI (colon) data; and 38% of HRRs had at least 50% of hospitals reporting CLABSI and MRSA data. SSI (hysterectomy) data were reported much less frequently, with only 4% of HRRs having >50% of hospitals reporting these data. As shown in Table 1, percentages were lower for the more stringent reporting thresholds (ie, higher percentages of hospitals reporting data).

Outcome 2: HAI performance diversity. Table 2 shows the percentage of HRRs where at least 1 pair of hospitals had statistically different SIRs. This is an indicator of the utility of HAI data within HRRs. If a HRR has at least 1 pair of hospitals with statistically different SIRs for a given HAI, then those data are useful for distinguishing between at least 2 hospitals within that HRR. CAUTI and *C. difficile* had the most diversity within HRRs, with 54% and 72% of HRRs having at least 1 pair of hospitals with statistically different SIRs for these HAIs, respectively. Lower percentages were observed for the other HAIs.

DISCUSSION

We analyzed publicly available HAI data to assess whether these data are (1) available among geographically proximate hospitals and (2) can be used to compare them.

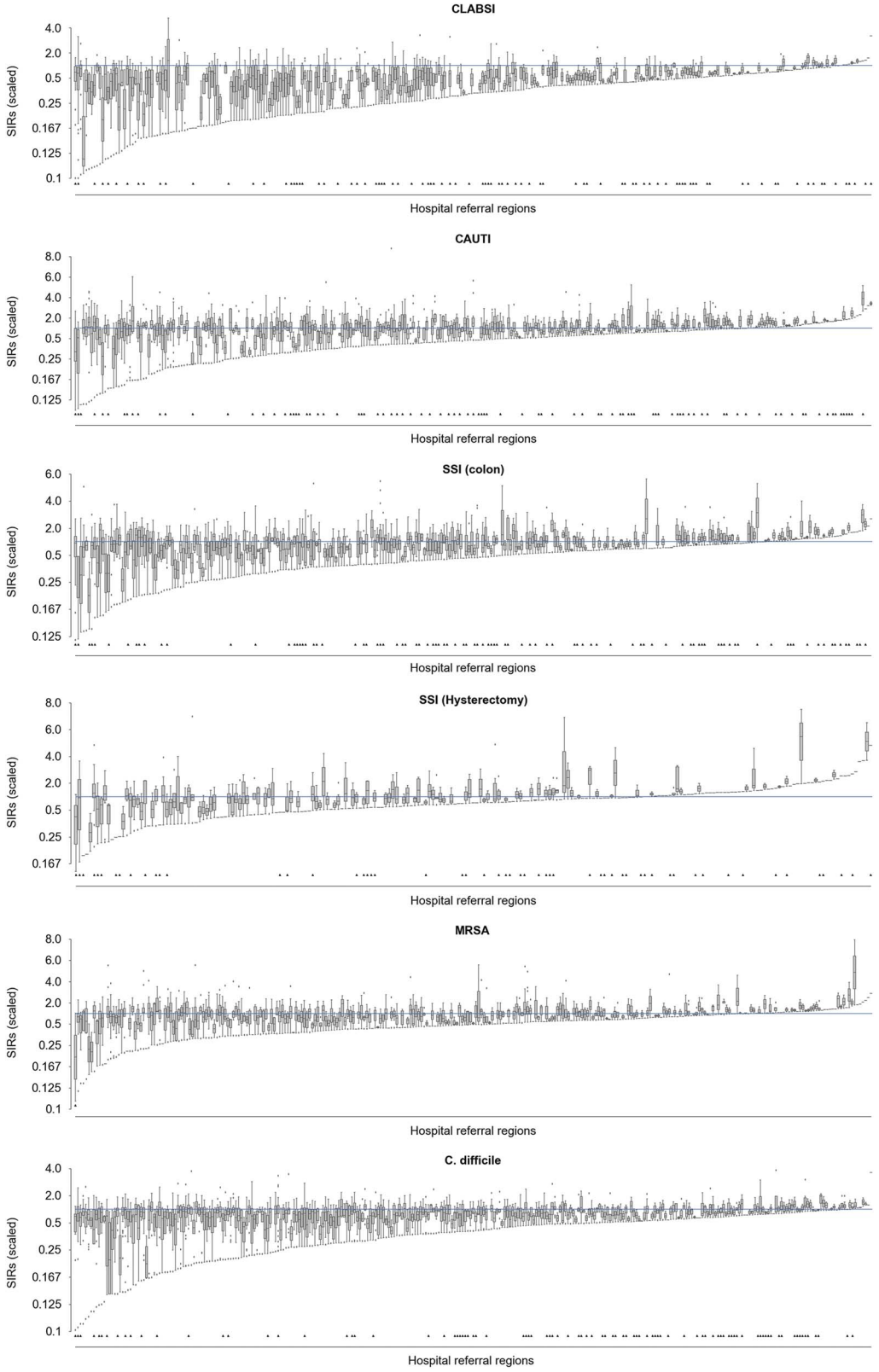

FIGURE 2. The distribution of standardized infection ratios (SIRs) within each hospital referral region (HRR), with a separate panel for each healthcare-associated infection. In each panel, a separate boxplot is shown for each HRR. The upward-pointing triangles (▲) below the boxplots indicate that a given HRR has one or more hospitals with SIR < 0.05. The vertical axis has been rescaled so that SIRs < 1 occupy the same amount of vertical space as SIRs > 1. The horizontal line indicates SIR = 1.

TABLE 1. Percentage of Hospital Referral Regions (HRRs) That Have at Least 50%, 70%, and 90% of Its Hospitals Reporting Data for Each Hospital-Associated Infection (HAI)

		>50% Reporting Data		>70% Reporting Data		>90% Reporting Data	
HAI	No. of HRRs	No.	%	No.	%	No.	%
CAUTI	306	144	47	92	30	47	15
Clostridium difficile	306	250	82	195	64	104	34
CLABSI	306	115	38	75	25	33	11
MRSA	306	117	38	65	21	25	8
SSI (hysterectomy)	306	12	4	1	0	0	0
SSI (colon)	306	131	43	67	22	24	8

NOTE. CAUTI, catheter-associated urinary tract infection; CLABSI, central-line–associated bloodstream infection; MRSA, methicillin-resistant *Staphylococcus aureus*; SSI, surgical site infection.

TABLE 2. Percentage of Hospital Referral Regions (HRRs) That Have at Least 1 Pair of Hospitals With a Statistically Significant ($P < .05$) Standardized Infection Ratio (SIR)

		HRRs with ≥1 Pair of Hospitals With Overlapping SIRs	
	Total No. of HRRs	No.	%
CAUTI	303	162	53.5
Clostridium difficile	306	220	71.9
CLABSI	303	99	32.7
MRSA	293	87	29.7
SSI (hysterectomy)	228	21	9.2
SSI (colon)	300	106	35.3

NOTE. CAUTI, catheter-associated urinary tract infection; CLABSI, central-line–associated bloodstream infection; MRSA, methicillin-resistant *Staphylococcus aureus;* SSI, surgical site infection.

In a substantial number of geographic regions (HRRs), >50% of the hospitals in each group report data (Outcome 1, data availability). *Clostridium difficile* had the highest percentage of reported data, with 82% of HRRs have at least 50% of hospitals reporting *C. difficile.* For CAUTI, 47% of HRRs have at least 50% of hospitals reporting this HAI, compared with SSI (colon) (43% of HRRs with at least 50% of hospitals reporting), CLABSI, and MRSA (38% of HRRs with at least 50% of hospitals reporting). Far fewer hospitals report SSI (hysterectomy) HAI data. Thus, HAI data generally are reported by enough hospitals for it to be useful in comparisons in many geographic locations, though this varies by HAI.

Outcome 2 assessed performance differences between hospitals within a geographic region. For each HAI, we calculated the percentage of hospitals in each geographic region where at least 1 pair of hospitals had statistically different SIRs. Results varied by HAI, ranging from 72% of geographic regions with at least 1 pair of hospitals with statistically different performance for *C. difficile*, to 9% for SSI (hysterectomy). These results indicate that CAUTI, *C. difficile*, and CLABSI data are more useful to differentiate hospitals than data for the other publicly reported HAIs.

To our knowledge, ours is the first in-depth assessment of its kind for publicly reported HAI data. Safavi et al[7] conducted a somewhat similar analysis for SSI process data and found that it was not useful in differentiating the "vast majority" of hospitals.[7] In contrast, our findings indicate that HAI outcome data are useful to differentiate some hospitals in many, though not all, geographic areas (corresponding to the areas covered by HRRs).

Strengths of this study include the use of nationwide data (as opposed to regional data) and including analysis of all 6 HAIs reported by the CMS. As described above, CMS HAI data are reported using SIRs, a method of indirect standardization to adjust for risk among hospitals. A major limitation of indirect standardization is that comparison of indirectly standardized rates such as SIRs can suffer from residual confounding (see Online Appendix for further discussion).[8]

Despite potential bias from residual confounding inherent in using SIRs, patients and other users of the CMS Hospital Compare website do compare SIRs between hospitals. This is the only available risk-adjusted HAI measure of hospital performance, and the primary purpose of the CMS Hospital Compare website is to compare hospitals. Outcome 2 of our analysis examines this usage of public HAI data, but it is limited by potential biases inherent in comparing SIRs among hospitals. The CMS does not publish the data necessary to determine the extent of these biases; it would be helpful if CMS made these data available.

In conclusion, HAI data currently available on the CMS Hospital Compare website are useful to differentiate hospitals. Unfortunately, the utility of these data is reduced by the number of hospitals reporting "not available" for HAIs (either the hospital does not report the data, or because of denominators below the CMS reporting threshold), and because comparing SIRs can be subject to bias. The current reporting of HAI data to the public via the CMS Hospital Compare website is a good first step, but it could be further improved with additional research on which outcome measures should be reported by hospitals to the CDC, with the addition of risk adjustment methodology for these measures.

ACKNOWLEDGMENTS

Financial support: This work was supported by the Agency for Healthcare Research and Quality (grant no. HS18111 to D.J.M.), the Veterans Administration Health Services Research and Development (grant no. CRE 12-289 to D.J.M.), the Baltimore VA Medical Center Geriatrics Research, Education, and Clinical Center (J.D.S.), the National Institute on Aging at the National Institutes of Health (grant no. NIA 5 P30 AG028747 to J.D.S.), the National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes of Health (grant no NIDDK 5 P30 DK072488 to J.D.S.), and the National Institutes of Health (grant no. 5 K24AI079040-05 to A.D.H.).

Potential conflicts of interest: The authors report no conflicts of interest. D.J.M. serves as a research consultant for Welch-Allyn, has presented a self-developed lecture for a 3M series on hospital infections, and receives travel support for conferences from IDSA, SHEA, and ASM.

Address correspondence to Max Masnick, 10 S. Pine Street, MSTF 330, Baltimore, MD 21201 (max@masnick.net).

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2017.179

REFERENCES

1. Healthcare-associated infection progress report: questions and answers. Centers for Disease Control and Prevention website. http://www.cdc.gov/HAI/surveillance/QA_stateSummary.html#b9. Updated 2016. Accessed July 28, 2017.
2. Rock C, Thom KA, Harris AD, et al. A multicenter longitudinal study of hospital-onset bacteremia: time for a new quality outcome measure? *Infect Control Hosp Epidemiol* 2016;37:143–148.
3. Healthcare-associated infections hospital data. US Medicare Services website. https://data.medicare.gov/Hospital-Compare/Healthcare-Associated-Infections-Hospital/77hc-ibv8/about. Accessed July 28, 2017.
4. Data by region. Dartmouth Atlas of Health Care website. http://www.dartmouthatlas.org/data/region/. Updated 2017. Accessed July 28, 2017.
5. Soe MM. Compare 2 SIRs by exact binomial test and mid-P, 95% CI. Centers for Disease Control and Prevention website. http://www.cdc.gov/nhsn/sas/binom.sas. Accessed October 24, 2016.
6. Fast Facts on US hospitals. American Hospital Association website. http://www.aha.org/research/rc/stat-studies/fast-facts.shtml. Published 2013. Accessed July 28, 2017.
7. Safavi KC, Dai F, Gilbertsen TA, Schonberger RB. Variation in surgical quality measure adherence within hospital referral regions: do publicly reported surgical quality measures distinguish among hospitals that patients are likely to compare? *Health Serv Res* 2014;49:1108–1120.
8. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2008.