Benchmarking local healthcare-associated infections: Available benchmarks and interpretation challenges

Aiman El-Saed\textsuperscript{a,b,c,d}, Hanan H. Balkhy\textsuperscript{a,b,d,e,*}, David J. Weber\textsuperscript{f}

\textsuperscript{a} Gulf Cooperation Council (GCC) States and WHO Collaborating Center for Infection Prevention & Control, Riyadh, Saudi Arabia
\textsuperscript{b} Department of Infection Prevention and Control, King Abdulaziz Medical City, Riyadh, Saudi Arabia
\textsuperscript{c} Community Medicine Department, Faculty of Medicine, Mansoura University, Egypt
\textsuperscript{d} King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
\textsuperscript{e} Department of Pediatrics, King Abdulaziz Medical City, Riyadh, Saudi Arabia
\textsuperscript{f} University of North Carolina Health Care, University of North Carolina, Chapel Hill, NC, USA

Received 2 February 2013; received in revised form 1 April 2013; accepted 6 May 2013

**KEYWORDS**
Benchmarking; Comparison; Surveillance; Healthcare-associated infections

**Summary**  Growing numbers of healthcare facilities are routinely collecting standardized data on healthcare-associated infection (HAI), which can be used not only to track internal performance but also to compare local data to national and international benchmarks. Benchmarking overall (crude) HAI surveillance metrics without accounting or adjusting for potential confounders can result in misleading conclusions. Methods commonly used to provide risk-adjusted metrics include multivariate logistic regression analysis, stratification, indirect standardization, and restrictions. The characteristics of recognized benchmarks worldwide, including the advantages and limitations are described. The choice of the right benchmark for the data from the Gulf Cooperation Council (GCC) states is challenging. The chosen benchmark should have similar data collection and presentation methods. Additionally, differences in surveillance environments including regulations should be taken into consideration when considering such a benchmark. The GCC center for infection control took some steps to unify HAI surveillance systems in the region. GCC hospitals still need to overcome legislative and logistic difficulties in sharing data to create their own benchmark. The availability of a regional GCC benchmark may better enable health care workers and researchers to obtain more accurate and realistic comparisons.

© 2013 King Saud Bin Abdulaziz University for Health Sciences. Published by Elsevier Ltd. All rights reserved.

* Corresponding author at: King Abdulaziz Medical City, Riyadh, Saudi Arabia. Tel.: +966 504623090.
E-mail address: balkhyh@hotmail.com (H.H. Balkhy).

1876-0341/$ – see front matter © 2013 King Saud Bin Abdulaziz University for Health Sciences. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jiph.2013.05.001
Introduction

Between 7 and 10% of patients worldwide admitted to acute care hospitals develop at least one healthcare-associated infection (HAI) during their hospital stay [1]. HAIs add extra morbidity and mortality risks to patients and lead to considerable stretching of many countries’ already limited healthcare resources [1–3]. Recently, HAI surveillance as part of a broad-based prevention and control strategy has received more attention from healthcare facilities, patient-safety organizations, and patients themselves [4]. Growing numbers of healthcare facilities are routinely collecting standardized data on HAIs, which are used not only to track internal performance but also to compare local data to national and international benchmarks [4].

Benchmarking

Prior to its use in healthcare surveillance, benchmarking was recognized in industry as an effective means of improving business performance [5]. Today, HAI benchmarking can be divided into internal and external systems. Internal benchmarking typically involves comparing current processes and/or outcomes to baseline data or comparing different departments in the same healthcare facility [6]. Although easily accessible and potentially highly useful, the collection of baseline data that is of adequate size for statistical comparison may require a significant amount of time. Moreover, the inability to adjust for patient, healthcare, and methodological changes over time may lead to erroneous conclusions. External benchmarking, on the other hand, usually involves comparing processes and/or outcomes in one healthcare facility to other facilities performing similar activities, often with higher standards [7]. The main challenge to external benchmarking is accounting for differences in patient risks and surveillance methodologies.

The purpose of both internal and external benchmarking is to continuously improve healthcare by demonstrating strengths and weaknesses, stimulating competitiveness, and assessing the value of interventions intended to reduce HAIs [6]. Benchmarking is often compromised by the limitation of simply comparing outcome indicators rather than analyzing and promoting the best practices [8]. Without performing these latter activities, the benchmarking of HAI data can be misleading. Furthermore, the benchmarked data must be collected using standardized case definitions as well as similar data collection methods and in populations of adequate sizes over a sufficient duration of time, as a statistically relevant number of outcomes are required for comparison [9]. Moreover, the collected data should be analyzed and reported using similar risk-stratified or risk-adjusted metrics (rates, proportions, or ratios) to allow fair comparisons [9]. Nevertheless, benchmarking is often performed without fulfilling these conditions, perhaps because local policy makers poorly understand the significance of these limitations. Obviously, external benchmarking cannot be accomplished if there is no regional system for data collection and dissemination.

Benchmarking risk-adjusted metrics

One of the major challenges in benchmarking metrics of HAI surveillance is the heterogeneity of healthcare facilities in terms of HAI risk. The potential for healthcare facilities to report higher rates of HAIs is dependent on many factors including size (bed number) of the facility, type and complexity of the care provided (such as burn care and solid organ transplants), length of patient stay, duration and type of device use, patient risks for an HAI (such as age and immunocompromising conditions), and comorbidities (such as renal dysfunction, liver failure, obesity, and diabetes) [10–13]. Therefore, benchmarking overall (crude) HAI surveillance metrics without accounting or adjusting for these variables can result in misleading conclusions. Providing risk-adjusted metrics is one way to reduce the possibility of such erroneous conclusions [4]. Statistical adjustments of the metric can take any of the following forms: (1) multivariate logistic regression analysis to adjust for multiple confounders at the same time; (2) stratification to adjust for (usually) one confounder at a time by stratifying the metric by the levels (groups) of that confounder; (3) standardization to adjust for (usually) one confounder based on weighted averages; or (4) restrictions to adjust for (usually) one confounder by excluding unwanted levels of that confounder.

Stratification is by far the most common adjustment method used in benchmark reports. The National Healthcare Safety Network (NHSN) and the International Nosocomial Infection Control Consortium (INICC) previously reported type-specific rates of device-associated HAI stratified by critical care unit types for adults and paediatric patients and by weight groups for neonatal patients [2,14]. Additionally, dialysis access-related infections were stratified according to the type of
vascular access [15], and procedure-specific surgical site infection (SSI) rates (actual proportions) were stratified according to the NHSN risk index category, which is based on the American Society of Anesthesiologists’ scores, procedure duration, and wound classification [16]. Although stratification is a straightforward and powerful method of adjustment, the question remains whether studies use the correct levels of stratification. For example, it was shown that procedure-specific stepwise logistic regression models for SSI data yielded new procedure-specific risk factors that were more predictive than the current risk index category [17]. Another potential problem with stratification is that as the rate of HAI decreases, small units (such as coronary care units) may have too few outcomes to allow statistically meaningful comparisons over a specified time (usually one month).

Multivariate regression adjustment and indirect standardization are increasingly used in reporting HAI surveillance metrics. A number of studies have adjusted HAI prevalence and antimicrobial use for the case-mix (i.e., heterogeneity regarding the patient’s risk) using multivariate logistic regression models and an indirect standardization method to allow for fair inter-hospital comparisons [11,18,19]. Approximately two decades ago, the National Nosocomial Infections Surveillance (NNIS) system introduced the standardized infection ratio (SIR) to indirectly standardize SSI rates using a standard population to enable fair comparisons of SSI rates between a healthcare facility and a benchmark with a different risk index category [20]. Recently, the NHSN promoted the expansion of SIR use to report a single SIR for a specified device-associated HAI from multiple hospital locations (such as specialty care areas) to adjust for differences in HAI incidence between these locations [21]. However, a recent report expressed doubts about the reliability and consistency of SIR compared to stratified HAI rates and showed that SIR may obscure, amplify, or reverse differences between two or more healthcare facilities and a benchmark due to its inherent dependence on the changes in risk-strata of both healthcare facilities and benchmark populations [22].

Benchmark reports

Benchmarks are typically public reports that apply a standard methodology and estimate risk-stratified or risk-adjusted HAIs and/or their preventive processes across a large network of healthcare facilities. Recognized benchmarks for HAI include the NHSN [23], INICC [24], European Centre for Disease Prevention and Control (ECDC), and World Health Organization (WHO) estimates [1]. The characteristics of these four benchmarks, including the advantages and limitations, are shown in Table 1.

1) NHSN reports: NHSN is a secure, internet-based surveillance system at the US Centers for Disease Control and Prevention (CDC) [23]. It was established in 2005 to integrate and replace three different surveillance systems at the CDC, including the NNIS, and NHSN is by far the most important and well-established surveillance system worldwide. One of its main stated purposes is to provide enrolled facilities with risk-adjusted metrics that can be used for inter-facility comparisons and local quality improvement activities. Starting in 2007, NHSN published a yearly report to estimate the magnitude of HAI, mainly in regards to risk-stratified pooled means and percentiles of device-associated and procedure-associated HAIs [14,16]. However, ignoring non-device-associated pneumonia, bloodstream infections, and urinary tract infections as well as some surgeries limits the comprehensiveness of the NHSN surveillance system [25]. The last antimicrobial resistance report was published by NNIS in 2004 [26], pointing to the frequency of reporting for some NHSN modules.

NHSN is widely used as a benchmark even outside of the US because its surveillance methodology is implemented in many hospitals worldwide. However, frequent changes in NHSN definitions, especially for catheter-associated urinary tract infection (CAUTI), dialysis events, antimicrobial use, and neonatal central line associated bloodstream infection (CLABSI), make it difficult for any healthcare facility outside the NHSN to interpret the results of their benchmarking if they do not incorporate these changes into their own surveillance system on a timely basis [27–29]. Approximately 90% of enrolled hospitals are general hospitals, including acute, trauma, and teaching facilities, although the number of enrolled hospitals has increased sharply during the last few years and now includes a larger representation of smaller hospitals.

2) INICC reports: The INICC is an international collaborative HAI surveillance system that uses a methodology largely similar to that of the NHSN [24]. It is the first international research network and largest of its type outside of the US. INICC was founded in Argentina in 1998, and the first multinational benchmark report
| **Table 1** Comparisons of the characteristics of recognized benchmarks. |
|---|---|---|---|---|
| **Covered countries** | NHSN | INICC | ECDC | WHO |
|  | US | 36 countries in South America, Asia, Africa, and Europe in 2009 | 17 European countries Up to 13 reported SSI | Systematic review of published data from 23 high- and 23 low-income countries |
| **Number of contributing hospitals** | Approximately 2500 in last report | 215 in last report | 1156 reported SSIs | Not defined |
| **Covered location for device-associated HAI** | ICU and non-ICU locations | ICU only | ICU only | ICU only |
| **HAI types covered** | SSI and device-associated HAI US CDC | Device-associated HAI Similar to US CDC Unit-based | SSI and device-associated HAI European CDC Unit-based and patient-based | SSI and device-associated HAI Mixed Mixed |
| **Type of device-associated HAI data** | Individual data are entered locally in an internet-based surveillance system and then centrally analyzed | Aggregate data are received from enrolled hospitals and then centrally analyzed | Individual data are entered in standardized national networks, and then data from all networks are centrally analyzed | Aggregate data are received from enrolled hospitals and then centrally analyzed |
| **Data entry & analysis** | Aggregate data are received from enrolled hospitals and then centrally analyzed | Aggregate data are received from enrolled hospitals and then centrally analyzed | Aggregate data are received from enrolled hospitals and then centrally analyzed | Aggregate data are received from enrolled hospitals and then centrally analyzed |
| **Advantages** | • Large data set that allows multiple stratifications • Uses standardized definitions of HAI • Reports device-associated HAI from ICU and non-ICU locations | • Covers under-studied limited-resource countries • Uses standardized definitions of HAI • Reports HAI-related mortality and length of stay as well as preventive bundles • Analyzes aggregate rather than individual data | • Large data set that allows stratifications and adjustments • Collects both unit-based and patient-based data • Provides some data adjusted for patient risk • Electronic data entry | • Good crude estimates for HAI incidence, prevalence, and impact • Covers both low- and high-income countries |
| **Limitations** | • Reports data on dialysis infections and antimicrobial use are infrequently released • No adjustment for patient risk • Not a true cohort, which epidemiologically limits comparing data over time | • No standardized electronic data collection in enrolled hospitals • No single-year data to examine changes over time • Hospitals included may not reflect their respective countries • Lack of device-associated HAI from non-ICU locations • Currently no SSI reports • No adjustment for patient risk • Validity of reported data is not determined | • Although standardized, definitions of HAI are not followed by all member countries • ECDC definitions are not popular outside of European countries • Frequent changes in surveillance systems over the last 2 decades limits the frequency of reports | • Includes studies with heterogenous case definitions of HAI and methods • Data presented are neither risk-stratified nor risk-adjusted • Data from low-income countries are fragmented and may not represent low-income countries |
(released in 2006) reported largely South American data [30]. Unlike NHSN, which receives locally entered individual data, INICC uses aggregate data received from enrolled hospitals that are sum and analyzed centrally [31]. Additionally, data on device-associated HAIs are collected from patients who do or do not develop HAI, allowing for comparisons of mortality and length of stay. Voluntary enrollment in INICC has dramatically increased over the last few years, and the 2012 report presented data from 215 hospitals from 36 countries in South America, Asia, Africa, and Europe [2]. In addition, almost all of the enrolled countries are considered low-resources countries, and the enrolled hospitals include academic teaching (44%), public (35%), and private community (21%) facilities and likely represent the elite hospitals in their respective countries. Similar to the NHSN, the INICC reports show the mean and percentiles of device-associated HAI rates and device utilization ratios as well as the rates of antimicrobial resistance. Additionally, these reports show data on related mortality, length of stay, and hand hygiene.

(3) ECDC reports: The ECDC is the current official HAI surveillance system in the European Union (EU) member states. ECDC uses the Hospitals in Europe Link for Infection Control through Surveillance (HELICS) protocol for SSI and ICU-acquired infections [32]. The main objective of the ECDC is to create an EU benchmark for inter-hospital comparisons of HAI rates, microorganisms, and antimicrobial resistance. In 2012 reports, between 11 and 13 out of 28 EU member states contributed data for device-associated HAI and SSI, while fewer states (n = 6) contributed data for antimicrobial use [33,34]. The HELICS protocol for SSI is quite similar to the NHSN protocol, and the SSI rate is stratified using the NHSN risk-index category. Unlike for the NHSN, the SSI rate in ECDC reports is additionally expressed per 1000 patient-days, and Poisson regression is used to adjust for case-mix defined according to the risk-index category. The HELICS protocol for ICU-acquired device-associated HAI is quite different from that of NHSN, as it collects both unit-based data (denominator) and patient-based data (risk factors) in patients who stayed three or more days in the ICU. Additionally, case definitions, particularly bloodstream infections and pneumonia, differ from those of the NHSN [35]. Unlike both NNIS and INICC reports that use

Figure 1 Device-associated HAIs in adult medical-surgical ICUs* of NHSN enrolled hospitals. *Data from different types of adult medical surgical ICUs (major teaching, ≤15 beds, and >15 beds) were combined when stratified.

defined daily use of antimicrobials data [26,36], the ECDC antimicrobial use data are presented as the number of treatment days for a specific antimicrobial per 1000 patient-days, and all ICUs are typically presented together [33].

(4) WHO HAI estimates: The WHO HAI estimates are based on systematic reviews published for national and multi-national data from both high- and low-resources countries, including the three benchmarks mentioned previously [1]. Because only a minority (15.6%) of low-income countries have published data on HAI incidence, this report likely does not reflect the actual HAI situation in low-income countries. Moreover, the data presented are neither risk-stratified nor risk-adjusted. However, the report can be viewed as a reasonable, albeit crude, source of estimates for HAI incidence, prevalence, and impact.

Choosing the right benchmark

The recent availability of benchmark reports from different parts of the world has widened the benchmarking options for new hospitals in GCC states. A chosen benchmark should have similar data collection and presentation methods, although the selection of the right benchmark report is not an easy task, particularly when there are wide variations in HAI incidence between benchmark reports using similar methods (Tables 2 and 3).
For example, the incidence of device-associated HAI is two- to three-fold higher in low-resources countries than in high-resources countries (Table 2). Additionally, even within high-resource countries, the incidence of SSIs is considerably different, as the incidence is lower in the NHSN than the ECDC for many procedures (Table 3). Moreover, the change in HAI incidence in consecutive reports from the same benchmark organization (Fig. 1) and the underlying contributing causes may complicate the selection and interpretation of the benchmarking process [14,16,26,37–42]. For example, several causes that may affect fair comparisons were hypothesized to explain the downward trend in device-associated HAI rates in consecutive NHSN reports, including (1) changes in HAI definitions to reduce the percentage of non-objective diagnoses (e.g., abandoning clinical sepsis as an acceptable diagnosis for CALBSI); (2) complying with regulations for mandatory HAI reporting in many states (this represented 70% of contributing hospitals in the 2010 data); (3) enrollment of many hospitals with smaller bed numbers, which generally have a lower risk of HAIs (this represented two-thirds of contributing hospitals in the 2010 data); and (4) implementation of multiple infection control strategies by many hospitals, which may have resulted in an actual decrease in HAI incidence.

### Table 2 Devic-associated HAIs and device utilization in adult medical-surgical ICUs in recognized benchmark reports.

<table>
<thead>
<tr>
<th></th>
<th>CA-BSI</th>
<th>CLU</th>
<th>CA-UTI</th>
<th>UCU</th>
<th>VAP</th>
<th>VU</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHSN (2010)</td>
<td>1.1 (1.1–1.2)</td>
<td>0.45</td>
<td>1.5 (1.4–1.5)</td>
<td>0.68</td>
<td>1.3 (1.2–1.4)</td>
<td>0.32</td>
</tr>
<tr>
<td>INICC (2004–2009)</td>
<td>5.9 (5.7–6.2)</td>
<td>0.53</td>
<td>7.1 (6.9–7.3)</td>
<td>0.56</td>
<td>18.4 (17.9–18.8)</td>
<td>0.38</td>
</tr>
<tr>
<td>ECDC (2007)</td>
<td>3.2</td>
<td>0.69</td>
<td>6.5</td>
<td>0.77</td>
<td>13.4</td>
<td>0.54</td>
</tr>
<tr>
<td>WHO, High-resource countries (1995–2010)</td>
<td>3.5 (2.8–4.1)</td>
<td>NA</td>
<td>4.1 (3.7–4.6)</td>
<td>NA</td>
<td>7.9 (5.7–10.1)</td>
<td>NA</td>
</tr>
<tr>
<td>WHO, Low-resource countries (1995–2010)</td>
<td>12.2 (10.5–13.9)</td>
<td>NA</td>
<td>8.8 (7.4–10.3)</td>
<td>NA</td>
<td>23.9 (20.7–27.1)</td>
<td>NA</td>
</tr>
</tbody>
</table>

- **CA-BSI**: Central line-associated, rather than catheter, in the NHSN and INICC reports. We excluded clinical sepsis from the INICC rate.
- **ECDC rates** included primary and secondary BSIs.
- **WHO estimates** were from all types of adult ICUs and included both catheter-related and -associated BSIs and UTIs.

### Table 3 Incidence of SSIs per 100 operations for selected operations in the NHSN (2006–2008) and ECDC (2008–2009).

<table>
<thead>
<tr>
<th></th>
<th>Risk index category</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Coronary artery bypass graft</td>
<td>NHSN</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>2.6</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>NHSN</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>0.9</td>
</tr>
<tr>
<td>Colon surgery</td>
<td>NHSN</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>7.2</td>
</tr>
<tr>
<td>Caesarean section</td>
<td>NHSN</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>3.5</td>
</tr>
<tr>
<td>Hip prosthesis</td>
<td>NHSN</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>0.7</td>
</tr>
<tr>
<td>Knee prosthesis</td>
<td>NHSN</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>0.6</td>
</tr>
<tr>
<td>Laminectomy</td>
<td>NHSN</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>ECDC</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### Benchmarking local data

Benchmarking local GCC data is challenging, although benchmarking to NHSN reports is preferred because the case definitions and methodologies are similar and differences in HAI rates will likely encourage improvements. However, differences in surveillance environments, including
regulations in GCC and NHSN hospitals, should be taken into consideration. Additionally, delays in implementing frequent NHSN changes in case definitions and methodologies could further complicate interpretation of the data. Benchmarking to INICC seems legitimate because of similar methodologies and challenges, as well as the availability of unique data on mortality, length of stay, and prevention. However, the use of aggregate data from enrolled hospitals does not account for the variability in surveillance adjudication between and within participating countries. Moreover, the benchmarking process is expected to improve infection control practices when using a benchmark of a higher standard. ECDC may be an alternative benchmark to GCC hospitals for SSIs and antimicrobial use and resistance. However, the considerable differences in device-associated HAI definitions likely limit its use as a benchmark for that purpose. WHO estimates for high-resource countries are driven by NHSN and ECDC data, while the estimates for low-resource countries are largely fragmented and not derived from a clear source. Additionally, failure to account for the wide variability in surveillance methods implemented in different parts of the world, as well as failure to risk-stratify different metrics of HAI, limits the benefit of the WHO estimates as a benchmark. The GCC center for infection control distributed its second edition of the GCC surveillance manual in 2011 and has conducted many surveillance training activities to unify HAI surveillance systems in the region. However, GCC hospitals still need to overcome legislative and logistic difficulties in sharing data to create their own benchmark. The availability of a regional GCC benchmark that addresses many of the above challenges may better enable health care workers and researchers to obtain more accurate and realistic comparisons and may positively impact infection control standards and patient safety in the region.

Funding
No funding sources.

Competing interests
None declared.

Ethical approval
Not required.

References


[22] Birnbaum D, Zarate R, Marfin A. SIR, you’ve led me astray! Infection Control and Hospital Epidemiology 2011;32:276–82.


