Technical Documentation

Cost of Healthcare-Associated Infections Calculator

February 9, 2011
1 Motivation

Healthcare-associated Infections (HAI)

Healthcare-associated infections (HAIs) continue to cause serious harm and even death for patients. [1-4] Each year, two million patients admitted to acute-care hospitals in the United States acquire an HAI. [5, 6] This results in approximately 90,000 deaths and increased patient care costs of $4.5 billion - $5.7 billion annually. [7] A recent Office of the Inspector General report on adverse events estimates that 35 percent of patients suffer from an HAI. [8] This issue has been recognized as a Tier One Priority by HHS in the recent Patient Protection and Affordable Care Act. [9] In addition, several of the HAIs are hospital-acquired conditions (HACs). The six HAIs are catheter-associated urinary tract infection (CAUTI); central line-associated bloodstream infection (CLABSI); multidrug-resistant organisms (MDRO), which are Clostridium difficile infection (CDI) and methicillin-resistant Staphylococcus aureus (MRSA); surgical-site infection (SSI); and ventilator-associated pneumonia (VAP). The overlapping HACs are CAUTI; vascular catheter-associated infection; and surgical-site infection, specifically mediastinitis following coronary artery bypass graft, SSI following certain orthopedic procedures, and SSI following bariatric surgery for obesity.

The Need for an HAI Calculator

Published literature provides robust point estimates that are well grounded in scientific observational methods. These have been used to define the magnitude of the HAI problem and drive change. Unfortunately, there are limitations due to the scope of the study, characteristic of the organizations where the study was performed, which make generalization to other organizations difficult. In addition, cost to undertake those studies make the updates expensive and infrequent. With the cost of care increasing nearly every year, the actual cost in today’s numbers may be quite different. (U.S. healthcare spending growth decelerated in 2009, increasing 4.0 percent compared to 4.7 percent in 2008. Total health expenditures reached $2.5 trillion, which translates to $8,086 per person or 17.6 percent of the nation’s Gross Domestic Product, up from 16.6 percent in 2008.) [10]

This HAI calculator was designed to assist organizations in better understanding the excess costs associated with HAIs and the value of preventing them. It will help hospital leaders better understand and/or validate their internal HAI cost estimates with relevant estimates of HAI
impact to complement data from published HAI sources. It is a predictive software model that provides more targeted estimates of HAI incidence, cost, and length-of-stay impact based on hospital size, geographic location, teaching status, and patient admissions.

The calculator is designed with two types of users in mind:

- Limited data access: Many infection prevention departments do not readily have access to pertinent data on the frequency, cost, and length of stay (LOS) associated with infections within their organization. This calculator will provide estimates of the cost and LOS associated with common HAIs based on a peer group of similar hospitals. The calculator also provides default patient volume values for organizations that do not have those data readily available. The default values are based on published articles and reports. [11-26]

- Benchmarking: For those organizations with access to the costs and LOS associated with infections, this calculator will provide helpful benchmark data.

For both scenarios, this tool will give you and your leadership a more accurate projection for making decisions about resource-budgeting needs for your infection prevention department, and about HAI preventive interventions in general.

2 HAI Calculator Development

Underlying the HAI Cost Calculator is a series of regression models. This section describes the data source, the codes used to define the infections, the populations examined for each infection, and the structure of the regression models.

Data

Data for the model came from AHRQ’s Healthcare Cost and Utilization Project (HCUP) National Inpatient Sample (NIS) (http://www.hcup-us.ahrq.gov/nisoverview.jsp). The HCUP NIS is the largest all-payer inpatient database in the U.S., capturing nearly 90% of inpatient discharges. The first version of the HAI calculator used data from the year 2007, which includes all inpatient discharges from 1,044 hospitals in 40 states. The full sample of adult (18+) medical and surgical discharges was
6,639,401. The calculator was recently updated using 2008 data and allowed.

The HCUP NIS contains discharge data included in the UB-92 discharge form, including patient demographics (gender, age, race), admission and discharge status, socioeconomic status (median income for ZIP Code), primary and secondary diagnosis codes, primary and secondary procedure codes, payer type, hospital characteristics (ownership, size, teaching status, cost-to-charge ratios), total charges, and total length of stay. Cost data were obtained by applying the hospital-specific ratio of cost-to-charges to charges for the admission.

The hospital size segmentation is defined according to region and teaching status, according to table 2 below.

Infections

We used International Classification of Diseases, Ninth Revision, Clinical modification (ICD-9) diagnosis codes to identify six specific HAIs: surgical-site infections (SSI), ventilator-associated pneumonia (VAP), urinary tract infections (UTI), central line-associated bloodstream infections (CLABSI), methicillin-resistant *Staphylococcus aureus* (MRSA), and *Clostridium difficile* (*C. diff.*). Specific codes used to define these infections are presented in Table 1, along with definitions of the total populations. Note that the analyses only apply to adults 18 and older.

Following de Lissovoy et al., we used a secondary ICD-9 diagnosis code of 998.59 (Other postoperative infection) to identify SSI [5]. This analysis was limited to patients with a surgical DRG.

Following AHRQ guidelines, we identified VAP using a secondary ICD-9 diagnosis code of 997.3 (Respiratory complications not elsewhere classified). This analysis was limited to patients with an ICD-9 procedure code indicating use of a ventilator (96.70, 96.71, and 96.72).

Table 1 reports the ICD-9 codes used to identify UTI. For the population we could not use billing codes to identify use of a urinary catheter. Therefore, we limited the analysis to just patients with a surgical DRG, since most surgery patients receive a urinary catheter.

To identify CLABSI, we used the standard used for AHRQ patient safety indicators. This identifies CLABSI using a secondary ICD-9 code of
999.31. The population includes all patients with a surgical DRG, with additional exclusions. Patients with a principal ICD-9 diagnosis code of 999.31 (Infection due to central venous catheter), 999.3 (Other infections), or 996.62 (Infection due to vascular device, implant and graft) were excluded from the population. In addition, patients with any DRG code for an immuno-compromised state or for cancer were excluded from the population.

MRSA was identified using a secondary ICD-9 diagnosis code of 041.12 (Methicillin-resistant *Staphylococcus aureus*). Note that this code does not distinguish between MRSA infection and MRSA colonization; therefore these results may be conservative. The population included all medical and surgical patients.

*C. diff.* was identified using a secondary ICD-9 diagnosis code of 008.45 (*Clostridium difficile*). This code does not distinguish between *C. diff.*-associated disease and *C. diff.* colonization, nor does it distinguish between toxigenic variants of the *C. diff.* bacteria and non-toxic varieties. It only indicates presence of the organism. The population includes all medical and surgical patients.

**Regression Models**

A series of linear regression models was fitted to the data in order to obtain estimates of the excess cost and LOS of each type of infection at each class of hospital. A team of collaborators evaluated the different statistical methods to determine which would provide the best fit for the data. The following were tested and presented in the table below: Unadjusted, Regression, Propensity Score Neighbor, Propensity Score Kernel, and Regression with DRG for SSI. This was repeated for the other five HAIs.
The linear regression model selected was the Regression DRG, as it was felt by the expert group that it provided the most conservative estimate and accounted for patient variation with DRG. The models were of the following form:

$$y_i = \beta_0 + \beta_1 I_i + x'\gamma + d'\delta + c'\theta + \varepsilon_i$$  \hspace{1cm} (1)$$

where $i$ indexes patients, $y_i$ is the outcome (either cost or LOS) for patient $i$; $\beta_0$ is an intercept term; and $\beta_1$ is the coefficient on the infection indicator $I_i$. We define $x$ as a vector of demographic variables (including age and gender) and $\gamma$ as its corresponding vector of coefficients; $d$ is a vector of indicators for clusters of DRGs and $\delta$ is its corresponding vector of coefficients; $c$ is a vector of AHRQ comorbidities and $\theta$ is its corresponding vector of coefficients. Finally, $\varepsilon_i$ is a zero-mean, (iid) normally distributed error term.

Separate regressions were fitted for each class of hospital, characterized by a unique region, setting, teaching status, and size. For example, a single regression of cost of SSI was fit for large, Northeast, urban, teaching hospitals, etc. The coefficient $\beta_1$ was taken as the estimate of the excess cost of the infection. Hospital size was defined using AHRQ criteria, which vary by region and setting (Table 2).

<table>
<thead>
<tr>
<th>Location</th>
<th>Size</th>
<th>Type</th>
<th>Unadjusted</th>
<th>Regression DRG</th>
<th>P-score Neighbor</th>
<th>P-score Kernel</th>
<th>Regression DRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midwest</td>
<td>Large</td>
<td>Non-teaching</td>
<td>30,049</td>
<td>28,695</td>
<td>26,917</td>
<td>26,383</td>
<td>20,820</td>
</tr>
<tr>
<td>Midwest</td>
<td>Large</td>
<td>Teaching</td>
<td>54,465</td>
<td>48,194</td>
<td>41,350</td>
<td>27,273</td>
<td>29,754</td>
</tr>
<tr>
<td>Midwest</td>
<td>Medium</td>
<td>Non-teaching</td>
<td>32,156</td>
<td>30,101</td>
<td>30,781</td>
<td>31,385</td>
<td>21,172</td>
</tr>
<tr>
<td>Midwest</td>
<td>Medium</td>
<td>Teaching</td>
<td>27,430</td>
<td>24,727</td>
<td>24,522</td>
<td>25,155</td>
<td>18,248</td>
</tr>
<tr>
<td>Midwest</td>
<td>Small</td>
<td>Non-teaching</td>
<td>17,486</td>
<td>16,119</td>
<td>17,403</td>
<td>16,077</td>
<td>12,667</td>
</tr>
<tr>
<td>Midwest</td>
<td>Small</td>
<td>Teaching</td>
<td>44,461</td>
<td>42,176</td>
<td>43,619</td>
<td>46,012</td>
<td>24,168</td>
</tr>
<tr>
<td>Northeast</td>
<td>Large</td>
<td>Non-teaching</td>
<td>24,661</td>
<td>22,307</td>
<td>20,900</td>
<td>16,756</td>
<td>14,401</td>
</tr>
<tr>
<td>Northeast</td>
<td>Large</td>
<td>Teaching</td>
<td>68,586</td>
<td>61,229</td>
<td>50,769</td>
<td>27,115</td>
<td>30,836</td>
</tr>
<tr>
<td>Northeast</td>
<td>Medium</td>
<td>Non-teaching</td>
<td>39,223</td>
<td>36,199</td>
<td>35,252</td>
<td>21,172</td>
<td>20,270</td>
</tr>
<tr>
<td>Northeast</td>
<td>Medium</td>
<td>Teaching</td>
<td>39,552</td>
<td>36,543</td>
<td>32,788</td>
<td>16,067</td>
<td>22,603</td>
</tr>
<tr>
<td>Northeast</td>
<td>Small</td>
<td>Non-teaching</td>
<td>15,849</td>
<td>15,042</td>
<td>14,828</td>
<td>18,081</td>
<td>8,657</td>
</tr>
<tr>
<td>Northeast</td>
<td>Small</td>
<td>Teaching</td>
<td>19,628</td>
<td>16,763</td>
<td>16,688</td>
<td>16,826</td>
<td>7,729</td>
</tr>
<tr>
<td>South</td>
<td>Large</td>
<td>Non-teaching</td>
<td>26,590</td>
<td>24,290</td>
<td>16,768</td>
<td>5,351</td>
<td>10,470</td>
</tr>
<tr>
<td>South</td>
<td>Large</td>
<td>Teaching</td>
<td>47,842</td>
<td>44,229</td>
<td>31,354</td>
<td>2,491</td>
<td>20,709</td>
</tr>
<tr>
<td>South</td>
<td>Medium</td>
<td>Non-teaching</td>
<td>26,486</td>
<td>24,415</td>
<td>23,735</td>
<td>20,146</td>
<td>16,558</td>
</tr>
<tr>
<td>South</td>
<td>Medium</td>
<td>Teaching</td>
<td>40,890</td>
<td>37,969</td>
<td>31,657</td>
<td>18,124</td>
<td>19,760</td>
</tr>
<tr>
<td>South</td>
<td>Small</td>
<td>Non-teaching</td>
<td>23,170</td>
<td>21,097</td>
<td>19,841</td>
<td>7,537</td>
<td>12,165</td>
</tr>
<tr>
<td>South</td>
<td>Small</td>
<td>Teaching</td>
<td>26,519</td>
<td>24,834</td>
<td>24,576</td>
<td>22,960</td>
<td>15,206</td>
</tr>
<tr>
<td>West</td>
<td>Large</td>
<td>Non-teaching</td>
<td>39,237</td>
<td>36,900</td>
<td>33,879</td>
<td>19,728</td>
<td>22,246</td>
</tr>
<tr>
<td>West</td>
<td>Large</td>
<td>Teaching</td>
<td>52,088</td>
<td>46,042</td>
<td>36,892</td>
<td>17,183</td>
<td>20,227</td>
</tr>
<tr>
<td>West</td>
<td>Medium</td>
<td>Non-teaching</td>
<td>26,643</td>
<td>24,408</td>
<td>24,310</td>
<td>21,469</td>
<td>16,114</td>
</tr>
<tr>
<td>West</td>
<td>Medium</td>
<td>Teaching</td>
<td>39,617</td>
<td>37,901</td>
<td>46,291</td>
<td>41,161</td>
<td>29,524</td>
</tr>
<tr>
<td>West</td>
<td>Small</td>
<td>Non-teaching</td>
<td>25,131</td>
<td>23,042</td>
<td>24,414</td>
<td>23,111</td>
<td>14,676</td>
</tr>
<tr>
<td>West</td>
<td>Small</td>
<td>Teaching</td>
<td>44,911</td>
<td>41,938</td>
<td>45,118</td>
<td>49,092</td>
<td>24,259</td>
</tr>
</tbody>
</table>

Average: 34,697  31,855  29,691  22,073  19,871
Minimum: 15,649  15,042  14,828  2,491  7,720
Maximum: 68,586  61,229  56,769  49,092  30,836
Results of the calculator were presented to the Greenlight Project collaborators (http://www.safetyleadersdev.org/greenlightMeeting/speakers.jsp) to validate the data.

3 HAI Calculator Instruction

This section will describe the calculator, how to navigate it, and how to interpret its output. The HAI Calculator in Excel format can be found here: http://safetyleaders.org/greenlightMeeting/HAImodule/CostCalculator12_30_10.xlsx.

Navigation

The Cost of HAI Calculator tool consists of basic worksheets. Worksheets are indicated by tabs at the bottom of the spreadsheet. The Instructions tab is a worksheet that contains instructions, and the Inputs tab is the worksheet where the user enters the institution’s characteristics. The remaining worksheets present results in a tabular or graphical format. You can click on each tab to view instructions, enter data, or view results as necessary. The intended sequence is to read the instructions, then enter data on the Inputs worksheet, and then view results on the Results worksheets.

Inputs

On the Inputs worksheet the user enters Region, Setting, and Teaching status by selecting the yellow cells and choosing a category from the drop-down menus. The user also enters the hospital’s total number of beds.

Region is defined according to the Agency for Healthcare Research and Quality (AHRQ). The country is partitioned into four large regions: Northeast, Midwest, South, and West. The Northeast includes ME, NH, VT, MA, RI, CT, NY, NJ, and PA. The Midwest includes OH, IN, IL, MI, WI, MN, IA, MO, ND, SD, NE, and KS. The South includes DE, MD, DC, VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, and TX. The West includes MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR, CA, AK, and HI.

Teaching Status is a dichotomous indicator indicating teaching hospitals and non-teaching hospitals. If a hospital has an AMA-approved residency program, is a member of the Council of Teaching Hospitals (COTH), or has a ratio of full-time equivalent interns and residents to beds of .25 or higher, then the hospital is considered a teaching hospital. Otherwise, it is considered a non-teaching hospital.
The calculator also provides default values for certain patient volumes that may not be available readily to certain users.

**Outputs**

Tables and graphs are generated from user input to summarize the total costs associated with infections and the total hospital days lost to infections. To view a table of these results, click on the Results worksheet Tabs. The Results worksheet presents the excess costs and LOS associated with each infection. It also provides the total costs and days lost, which is the product of the excess attributable cost and the number of infections expected at the peer group of hospitals. The method for producing these estimates will be described in the next section. Here we describe the remaining output worksheets.

**Expected Cost Percent**
This graph presents the expected percent of the total cost burden of HAIs attributable to each type of infection based on the model. This may be helpful in identifying the infections with the greatest cost burden.

**Actual Cost Percent**
This graph presents the actual percent of the total cost burden of HAIs attributable to each type of infection based on your numbers. This may be helpful in identifying the infections with the greatest cost burden.

**Expected LOS Percent**
This graph presents the expected percent of the total LOS burden of HAIs attributable to each type of infection based on the model. This may be helpful in identifying the infections with the greatest opportunity cost.

**Actual LOS Percent**
This graph presents the actual percent of the total LOS burden of HAIs attributable to each type of infection. This may be helpful in identifying the infections with the greatest opportunity cost.

**Costs Per HAI**
This graph presents the expected excess costs attributable to each infection. This shows the expected costs for a comparable peer group of hospitals.

**Costs Per HAI Variance**
This graph presents the expected excess costs attributable to each infection alongside the actual excess costs, if you entered them. This
provides a side-by-side comparison of your costs to those of a comparable peer group of hospitals. If you did not provide actual excess costs, then only expected costs are presented.

**Expected Total Excess Costs**
This graph presents the total excess costs of HAI. It is the product of the expected number of HAIs and the excess costs per HAI.

**Actual Total Excess Costs**
This graph presents the total excess costs of HAI. It is the product of the inputted number of HAIs and the excess costs per HAI.

**Total Cost Variance**
This graph presents the total excess costs attributable to each infection alongside the actual total excess costs, if you entered them. This provides a side-by-side comparison of your total costs to those of a comparable peer group of hospitals. If you did not provide actual excess costs, then only expected costs are presented.

**LOS Per HAI**
This graph presents the expected excess days attributable to each infection. This is the expected LOS for a comparable peer group of hospitals.

**LOS Per HAI Variance**
This graph presents the expected excess LOS attributable to each infection alongside the actual excess LOS, if you entered them. This provides a side-by-side comparison of your days lost to those of a comparable peer group of hospitals. If you did not provide actual excess LOS, then only expected days are presented.

**Total Excess LOS**
This graph presents the total excess days lost due to HAI. It is the product of the expected number of HAIs and the excess LOS per HAI.

**Total LOS Variance**
This graph presents the total excess days lost attributable to each infection alongside the actual total excess days, if you entered them. This provides a side-by-side comparison of your total days lost to those of a comparable peer group of hospitals. If you did not provide actual excess LOS, then only expected days are presented.
<table>
<thead>
<tr>
<th>Healthcare-Acquired Infection</th>
<th>Infection Code</th>
<th>Population</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical-Site Infection (SSI)</td>
<td>ICD-9 DX code 2-15 of 996.61 or 998.59</td>
<td>All surgical discharges, 18 years and older</td>
<td>de Lissovoy et al., <em>Am J Infect Control</em> 2009 Jun;37(5):387-97.</td>
</tr>
<tr>
<td>Urinary Tract Infection (UTI)</td>
<td>ICD-9 DX code 2-15 of 996.64, 111.2, 590.10, 590.11, 590.9, 590.80, 590.81, 595.0, 595.3, 595.4, 595.89, 595.9, 597.0, 597.8, 599.0</td>
<td>All surgical discharges, 18 years and older</td>
<td>Zhan et al., <em>Med Care</em> 2009 Mar;47(3):364-9.</td>
</tr>
<tr>
<td>Catheter-Related Bloodstream Infection (CRBSI)</td>
<td>ICD-9 DX code 2-15 of 999.31</td>
<td>All surgical and medical discharges, 18 years and older</td>
<td><a href="http://www.qualityindicators.ahrq.gov/TechnicalSpecs41.htm#PSI41">http://www.qualityindicators.ahrq.gov/TechnicalSpecs41.htm#PSI41</a></td>
</tr>
</tbody>
</table>

Exclude cases:
- with ICD-9-CM code of 999.3, 999.31, or 996.62 in the principal diagnosis field or secondary
- with length of stay less than 2 days
- with any code for immuno-compromised state (DRG: 10, 82, 203, 172, 274, 346, 489) or cancer (DRG: 102, 302, 480, 495, 481)

---

Table 1: Definitions of infections and populations
<table>
<thead>
<tr>
<th>REGION</th>
<th>SETTING</th>
<th>TEACHING</th>
<th>HOSPITAL BEDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast 1</td>
<td>Rural</td>
<td>Nonteaching</td>
<td>1-49 50-99 100+</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Teaching</td>
<td>1-124 125-199 200+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-249 250-424 425+</td>
</tr>
<tr>
<td>Midwest 2</td>
<td>Rural</td>
<td>Nonteaching</td>
<td>1-29 30-49 50+</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Teaching</td>
<td>1-74 75-174 175+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-249 250-374 375+</td>
</tr>
<tr>
<td>South 3</td>
<td>Rural</td>
<td>Nonteaching</td>
<td>1-39 40-74 75+</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Teaching</td>
<td>1-99 100-199 200+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-249 250-449 450+</td>
</tr>
<tr>
<td>West 4</td>
<td>Rural</td>
<td>Nonteaching</td>
<td>1-24 25-44 45+</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>Teaching</td>
<td>1-99 100-174 175+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-199 200-324 325+</td>
</tr>
</tbody>
</table>

**Table 2: Definitions of hospital size**
References


******************************************************************************


__________________________________________________________________________