National Center for Emerging and Zoonotic Infectious Diseases

Obtaining the number of predicted events for the Standardized Infection Ratio (SIR)

National Healthcare Safety Network Statistics Team

Educational Slide Deck

September 2024

Purpose

- The purpose of this slide deck is to provide users with a high-level summary of the methods used to generate the risk adjustment models for the 2022 national HAI baseline.
- Users can review the slides at their own pace
- More information about the risk adjustment models themselves, or about how the SIR or SUR metric are calculated and interpreted, can be found in NHSN's Guide to the 2022 Baseline SIRs: **https://www.cdc.gov/nhsn/2022rebaseline/analysis-resources.html**
- More information about the HAI Rebaseline process can be found on the Rebaseline homepage: **https://www.cdc.gov/nhsn/2022rebaseline/index.html**

Standardized Infection Ratio (SIR)

$$\text{SIR} = \frac{\#\ observed\ HAIs}{\#\ predicted\ HAIs}$$

HAIs reported to NHSN

Calculated by CDC

- When # of observed HAIs is greater than the # predicted, the SIR will be greater than 1

$$\frac{\text{5 observed CLABSIs}}{\text{3.2 predicted CLABSIs}} = \text{SIR of } \mathbf{1.6}$$

- If # observed HAIs is less than # predicted, the SIR will be less than 1
- P-values and 95% confidence intervals (CI) provide information about statistical significance
  - NHSN’s Statistics Calculator:
  - https://www.cdc.gov/nhsn/pdfs/ps-analysis-resources/statscalc.pdf
  - NHSN’s Guide to the 2022 Baseline SIRs: https://www.cdc.gov/nhsn/2022rebaseline/analysis-resources.html

How does CDC calculate the number of predicted events?

- **The number of predicted infections is calculated using a formula derived using statistical risk adjustment models**
  - CDC uses logistic regression for surgical site infections (SSI) and Negative Binomial Generalized Linear Models (GLMs) for most other types of HAI
- **Models use characteristics (factors) reported to NHSN that significantly impact HAI incidence**
  - Models estimated based on aggregate national HAI data from a baseline time period
- **Example: Critical Access Hospital MRSA Bacteremia LabID Event Model**

$$\begin{aligned} \text{Number of Events} = \ & \mathit{Exp}[-11.3451 + 0.9991 \times (\text{Outpatient CO prev. rate} > 0) \\ & + 0.8824 \times (\text{Inpatient CO prev. rate} > 0)\ ] \\ & \times (\text{MRSA Bacteremia Patient Days}) \end{aligned}$$

How does CDC estimate the formula for the SIRs? (1/2)

- **CDC obtains the formula (model) for the number of predicted events by estimating the parameters based on a single baseline year**
- **Facility data for a single (baseline) year are obtained from the NHSN database, according to the exclusion criteria for each HAI**
  - The dataset includes the characteristics (covariates) that will be assessed for the model
  - These characteristics may be continuous (numeric) or categorical variables
- **Each covariate is included in isolation in the (“univariate”) model to determine the optimal parameterization**
  - This includes discretization of continuous data into ordinal levels, as well as combination of ordinal or categorical levels

How does CDC estimate the formula for the SIRs? (2/2)

- **The final model was chosen by finding the optimal parameterizations of all covariates in linear regression procedures**
  - This final model is a linear combination of the optimal set of statistically significant covariates
  - Levels of covariates included in the optimal model that were not statistically significant were collapsed
- **Standard model diagnostics are used to ensure the assumptions of the technique are appropriately met**
  - Includes checks of residuals, goodness of fit, leverage and influence, collinearity
  - Model calibration and bootstrap validation were used to ensure predictive accuracy and generalization of model to new data

How do I interpret an SIR model?

Risk Factors Used in the Critical Access Hospital MRSA Bacteremia LabID Event Model

Factor	Parameter Estimate	P-Value
Intercept	-11.3451	<0.0001
Outpatient CO prevalence rate: > 0 per 100 encounters	0.9991	0.0003
Outpatient CO prevalence rate: 0 per 100 encounters or no applicable locations	Referent	-
Inpatient CO prevalence rate: > 0 per 100 admissions	0.8824	0.0098
Inpatient CO prevalence rate: 0 per 100 admissions	Referent	-

- **Table shows intercept and coefficients of the model with p-values**
  - Referent level was generally chosen as the lowest level for the 2022 baseline models
- **Model for predicted number of events:**

$$\text{Number of Events} = \mathit{Exp}[-11.3451 + 0.9991 \times (\text{Outpatient CO prev. rate} > 0) + 0.8824 \times (\text{Inpatient CO prev. rate} > 0)] \times (\text{MRSA Bacteremia Patient Days})$$

- **Example facility 1:** Outpatient CO prev. rate=0.015, Inpatient CO prev. rate=0.002

$$\text{Number of Events} = \mathit{Exp}\,[-11.3451 + 0.9991 + 0.8824] \times (\text{MRSA Bacteremia Patient Days})$$

- **Example facility 2:** Outpatient CO prev. rate=0.008, Inpatient CO prev. rate=0

$$\text{Number of Events} = \mathit{Exp}\,[-11.3451 + 0.9991 + 0] \times (\text{MRSA Bacteremia Patient Days})$$

- **Note that example facility 2 had Inpatient CO prev. rate=0, so referent level applied with zero coefficient**

For any questions or concerns, contact the NHSN Helpdesk

- **Use subject line: “2022 HAI Rebaseline”**
- **NHSN-ServiceNow** to submit questions to the NHSN Help Desk.
- Access new portal at **https://servicedesk.cdc.gov/nhsncsp**.
- If you do not have a SAMS login, or are unable to access ServiceNow, you can still email the NHSN Help Desk at nhsn@cdc.gov.

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 cdc.gov
Follow us on X (Twitter) @CDCgov & @CDCEnvironment