Understanding the 2022 NHSN Rebaseline

Andrea Parriott
California Department of Public Health HAI Program

Table of Contents

1   Understanding the SIR

2   Rebaseline process

3   Rebaseline rationale

4   How will this affect my facility's SIR?

5   Q&A

Acknowledgment

Many of the figures used in this presentation were created by the US Centers for Disease Control and Prevention.

CDPH

Confidential - Low

3

Understanding the SIR

SIR Formula

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

HAIs reported to NHSN

Calculated by CDC

Source: Centers for Disease Control and Prevention

CDPH

5

Confidential - Low

Where do predicted cases come from?

Predicted cases are calculated using formulas that incorporate *parameters* from *generalized linear models* fit using *baseline data*.

CDPH

Confidential - Low

Simple linear regression models

- y is variable we are trying to predict
- x is variable we are using to predict y
  - Explanatory variable
- $E(y) = \alpha + \beta x$
- α is the intercept parameter
- β is the slope parameter
- Everything right of the equal sign is called the linear predictor
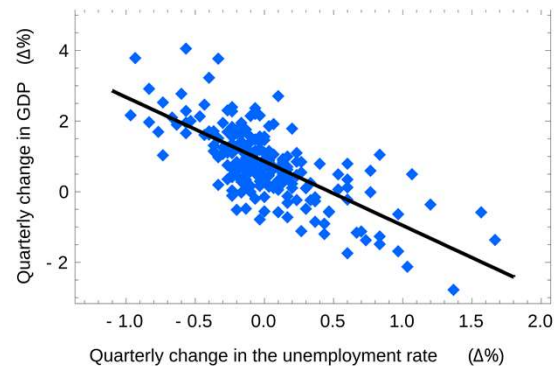
Source: Wikimedia Commons

CDPH

7

Confidential - Low

Multiple linear regression models

- More than one explanatory variable
- $E(y) = \alpha + \beta_1 x_1 \ldots + \beta_n x_n$
- α is value of y when all the explanatory variables are equal to zero
- β_k is the slope of the best fit line plotting y and x_k while holding all other x constant
- Categorical variables are coded as dummy variables

CDPH

8

Confidential - Low

Generalized linear regression models

- $E(y) = f(\alpha + \beta_1 x_1 \ldots + \beta_n x_n)$

- f is inverse link function

- Better fit for certain types of outcome data
  - Negative binomial regression predicts count variables or incidence rates
  - Logistic regression predicts probability of an event

CDPH

Confidential - Low

9

How predicted cases are calculated

- Baseline data is used to fit a generalized linear regression model
  - Baseline data is nationwide data from a single year

- Negative binomial regression models are used to predict count of infections in the LabID and Device Associated Modules

- Logistic regression models are used to predict the probability of infection in the Procedure Associated Module

- Regression parameters are used with facility/patient/procedure characteristics to get predicted cases

CDPH

Confidential - Low

10

How predicted cases are calculated

- LabID Module
  - Predicted cases are calculated at the facility level
- Device-associated Module
  - Predicted cases are estimated at the patient care area level
  - Facility predicted cases is sum of patient care area predicted cases
- Procedure associated module
  - Risk of infection is predicted for each procedure
  - Facility predicted cases is sum of procedure risks
- See *The NHSN Standardized Infection Ratio (SIR): Guide to the SIR* for more details

CDPH

Confidential - Low

11

Example: MRSA BSI in CAH

CAH MRSA BSI Model Parameters

Factor	Parameter Estimate
Intercept	-11.3451
Outpatient CO prevalence rate[1]: > 0 per 100 encounters	0.9991
Outpatient CO prevalence rate[1]: 0 per 100 encounters or no applicable locations	REFERENT
Inpatient CO prevalence rate[2]: > 0 per 100 admissions	0.8824
Inpatient CO prevalence rate[2]: 0 per 100 admissions	REFERENT

Source: Centers for Disease Control and Prevention

Predicted Cases Formula

Predicted cases =
$\exp[-11.3451$
$+0.9991(\textit{Outpatient CO prev.rate}>0)$
$+0.8824(\textit{Inpatient CO prev.rate}>0)]$
$\times \textit{patient days}$

CDPH

Confidential - Low

12

Example: MRSA BSI in CAH

CAH MRSA BSI Model Parameters   **Example**

Factor	Parameter Estimate
Intercept	*-11.3451*
Outpatient CO prevalence rate[1]: > 0 per 100 encounters	0.9991
Outpatient CO prevalence rate[1]: 0 per 100 encounters or no applicable locations	REFERENT
Inpatient CO prevalence rate[2]: > 0 per 100 admissions	0.8824
Inpatient CO prevalence rate[2]: 0 per 100 admissions	REFERENT

Source: Centers for Disease Control and Prevention

- Facility with:
  - 4,000 patient days
  - 1 outpatient CO MRSA BSI
  - 2 inpatient CO MRSA BSI

- Predicted cases =
  exp[-11.3451
  + 0.9991*1
  + 0.8824*1]*4,000
  = 0.3104

CDPH

Confidential - Low

13

Example: Cardiac surgery

2015 Cardiac Surgery Model Parameters   **Infection Risk Formula**

Factor	Parameter Estimate
Intercept	-5.2166
Emergency: Yes	0.7164
Emergency: No	REFERENT
Affiliated with a major teaching hospital: Yes	0.2867
Affiliated with a major teaching hospital: No	REFERENT
Age at time of procedure/10	-0.1021
Procedure duration: ≥236	0.3903
Procedure duration: <236	REFERENT
BMI: ≥30	0.6013
BMI: <30	REFERENT

- Predicted Risk =
  expit[-5.2166
  + 0.7164(*Emergency*)
  + 0.2867(*Teaching Affiliation*)
  - 0.1021(*Age/10*)
  + 0.3903(*Procedure Duration >= 236*)
  + 0.6013(*BMI>=30*)]

- expit(*x*) = $\frac{\exp(x)}{1+\mathrm{ex}\ (x)}$

CDPH

Confidential - Low

14

Example: Cardiac surgery

2015 Cardiac Surgery Model Parameters

Factor	Parameter Estimate
Intercept	-5.2166
Emergency: Yes	0.7164
Emergency: No	REFERENT
Affiliated with a major teaching hospital: Yes	0.2867
Affiliated with a major teaching hospital: No	REFERENT
Age at time of procedure/10	-0.1021
Procedure duration: ≥236	0.3903
Procedure duration: <236	REFERENT
BMI: ≥30	0.6013
BMI: <30	REFERENT

Infection Risk Formula

- 57 year old patient with BMI of 25, emergency surgery in major teaching hospital, duration=200

- Predicted Risk =
expit[-5.2166
+ 0.7164*1
+ 0.2867*1
- 0.1021*5.7
+ 0.3903*0
+ 0.6013*0] =
0.0082

CDPH

15

Confidential - Low

Example: Cardiac Surgery

Patient	Predicted Risk
1	0.0082
2	0.0065
3	0.0051
4	0.0033
5	0.0061
6	0.0067
7	0.0116
8	0.0047
9	0.0053
10	0.0163

Facility predicted cases=

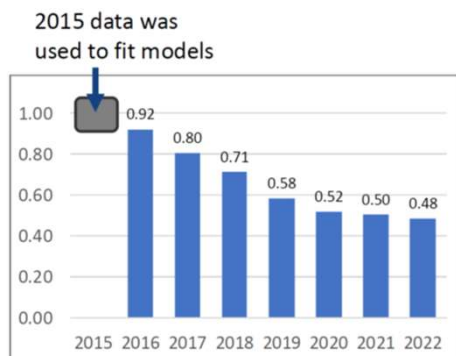0.0082 + 0.0065 + 0.0051 + 0.0033 +0.0061 + 0.0067 + 0.0116 + 0.0047 + 0.0053 + 0.0163 =

0.0738

CDPH

Confidential - Low

How is the new National Baseline created?

2015 National baseline: calibrated to 2015 data

2015 data was
used to fit models

- The model used to predict the number of events was fitted using 2015 data
  - The resulting predictions in any year using the 2015 baseline model are consistent with 2015

- For 2015, the mean SIR over all facilities would have an SIR of approximately 1
  - If observed events are reduced, in following years (and all other factors remain the same) the SIR is reduced

Source: Centers for Disease Control and Prevention

CDPH

Confidential - Low

9

2022 National baseline: calibrated to 2022 data

2022 data used to fit models

2015 2016 2017 2018 2019 2020 2021 2022
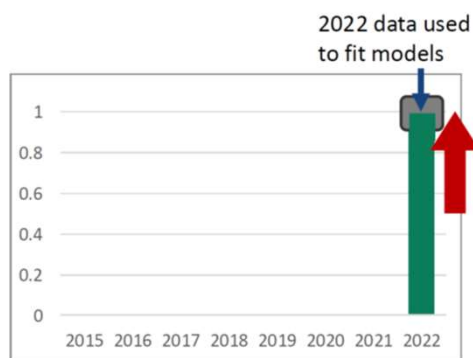
- The updated models use data from 2022 so resulting predictions in any other future year would be consistent with 2022

- SIRs calculated using the new 2022 National Baseline would therefore be recalibrated

- Average 2022 SIR across all facilities will be approximately equal to one

Source: Centers for Disease Control and Prevention

CDPH

Confidential - Low

Rebaseline process

- New regression models are created using 2022 national data

- Stepwise variable selection
  - Variables used in model may change

- Variables that were previously continuous may be changed to categorical and vice-versa

- Cutpoints for categories of continuous variables may change

- **Note: there are no substantial changes in definitions, exclusions or reporting**

CDPH

20

Confidential - Low

Progress on Rebaselined Model Release

SIR/SUR Reports (2022 Baseline) Available in NHSN	SIR/SUR Reports (2022 Baseline) Under Development in NHSN
- MRSA Blood LabID - CDI LabID - CAUTI - CLABSI and MBI-LCBI - VAE and pedVAE - SSI – Inpatient Adult Data - Complex 30-day SSI - All SSI - Complex Admission/Readmission SSI	- SSI – Inpatient Pediatric Data - All SSI - Complex Admission/Readmission SSI - SSI – Hospital Outpatient Procedures Data (HOPD) - Adult HOPD SSI - Pediatric HOPD SSI - TAP Reports/Dashboard - SUR Reports

Source: Centers for Disease Control and Prevention

CDPH

21

Confidential - Low

Why rebaseline?

"As a nation, we are getting better at surveillance and prevention of HAIs, and creating a new baseline, also called rebaseline, is a way to incorporate changes in detection practices and establish an updated national standard to continue prevention and reduction of HAIs."
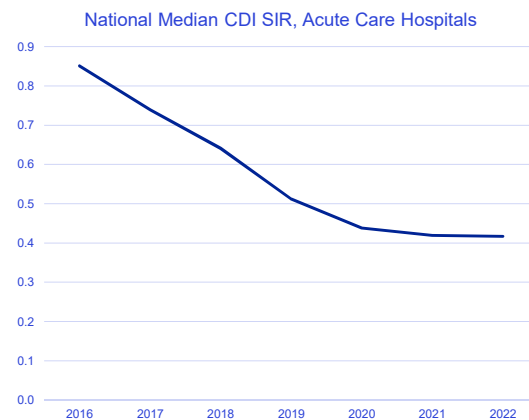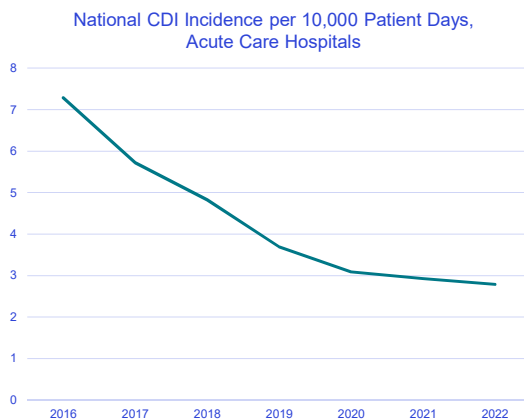
-CDC, What is the 2022 HAI Rebaseline and Why is it Important?

CDPH

23

Confidential - Low

Changes in Infection Prevalence

National CDI Incidence per 10,000 Patient Days, Acute Care Hospitals

National Median CDI SIR, Acute Care Hospitals

CDPH

24

Confidential - Low

Changes in Risk Factor Importance

2015 Baseline Colon Procedure

- Diabetes
- ASA score
- Sex
- Age
- BMI
- Closure technique
- Oncology hospital

2022 Baseline Colon Procedure

- Scope
- Procedure duration
- Trauma
- ASA score
- Age
- Surgical wound class
- Hospital bedsize
- Emergency procedure
- Medical school affiliation
- Sex
- BMI
- Oncology hospital
- General anesthesia

CDPH

Confidential - Low

25

How will this affect my facility's SIR?

Some SIRs will likely increase
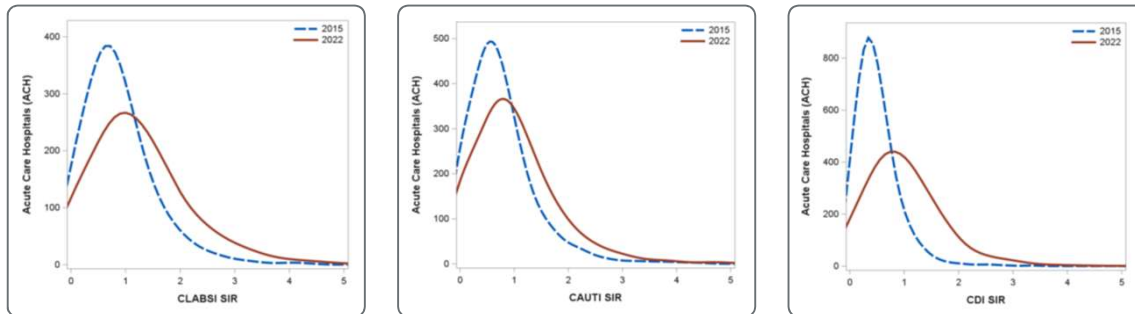
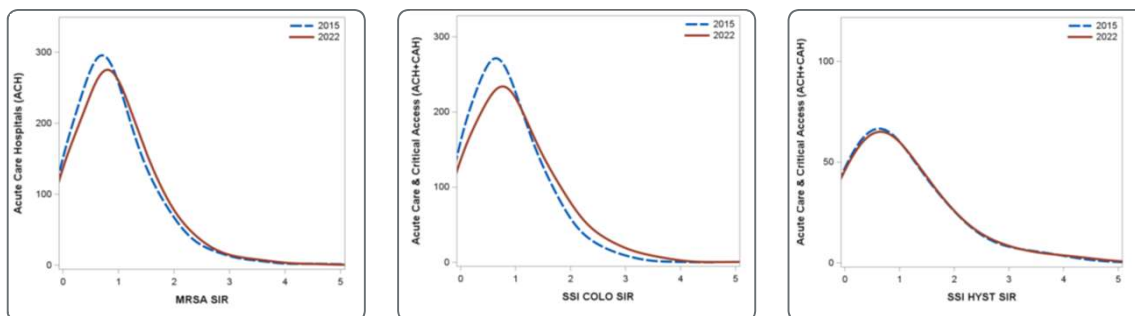**Most facility's SIRs will increase for infections where nationwide incidence rates have declined.**

Source: Centers for Disease Control and Prevention

CDPH

27

Confidential - Low

But not all…

Source: Centers for Disease Control and Prevention

CDPH

28

Confidential - Low

Rule 1

Do not compare 2015 baseline SIRs with 2022 rebaseline SIRs!

Comparisons between 2015 and 2022 baselines aren't meaningful

CDPH

Confidential - Low

29

Use one baseline to look at trends

- 2015 baseline files are still available for all years from 2015 to 2025.
- 2022 rebaselined files are available for 2022 to 2025.
- Use 2015 baseline for any interval that begins between 2015 and 2021

Source: Centers for Disease Control and Prevention

CDPH

Confidential - Low

30

Accessing 2022 and 2015 baseline reports

Source: Centers for Disease Control and Prevention

CDPH
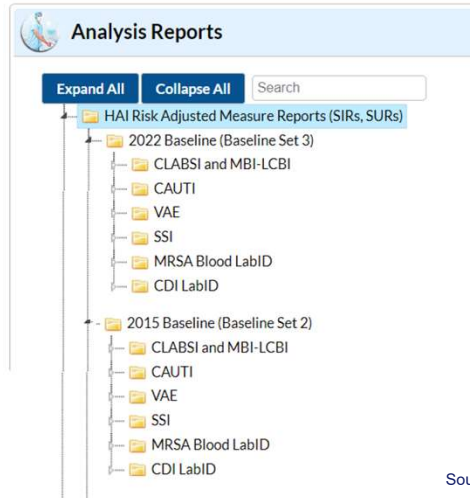
Confidential - Low

Summary

- Expected cases for SIRs are generated using formulas derived from regression model parameters

- These models are occasionally redone using more recent data to account for changes in incidence and risk factor importance

- SIRs from different baseline years shouldn't be compared

CDPH

Confidential - Low

32

Questions?

Andrea.Parriott@cdph.ca.gov or HAI_DATA@cdph.ca.gov

CDPH

Confidential - Low