Making the Most of Your Surveillance Data:
Biostatistics for Infection Control

Emily Sickbert-Bennett, PhD, MS, CIC, FSHEA

Director of Infection Prevention, UNC Medical Center
Professor of Medicine-Infectious Diseases, UNC School of Medicine
Associate Professor of Epidemiology, Gillings School of Global Public Health

1

Statistics

The margin of error...

17 in every 100 people...

Men are at 3 times higher risk...

Numbers that describe the health of the population

1 in 9 children...

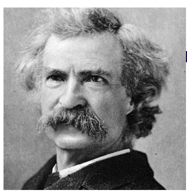
39% OF THE POPULATION...

The **science** used to interpret these numbers.

There is a statistically significant difference...

Risk of dying is 8 times higher among...

2



"There are 3 kinds of lies. Lies, damned lies, and statistics."

~Popularized by Mark Twain

- Describes the persuasive power of numbers, particularly the use of statistics, to bolster weak arguments, and the tendency of people to disparage statistics that do not support their positions.

3

Outline

- Describe Surveillance Data
- Display and Interpret Surveillance Data
- Determine the Significance of Changes to Surveillance Data

4

Describing Surveillance Data Using Descriptive Statistics

5

Absolute Measures

- Simplest type of measurement
- Also known as counts
- Example:
  - Hospital A: 25 patients with norovirus
  - Hospital B: 10 patients with norovirus
- Is norovirus worse at Hospital A?

6

Relative Measures

- Includes a denominator
- Useful for comparisons
- Examples:
  - 16 cases of *C. difficile* out of 1000 patients
  - 1 positive *C. difficile* test out of 7 samples tested

7

Absolute versus Relative

Example: Norovirus activity at different hospitals

- Absolute measures
  - Hospital A: 25 patients ill
  - Hospital B: 10 patients ill
- Relative measures
  - Hospital A: 25 ill per 1000 patients = 0.025 or 2.5%
  - Hospital B: 10 ill per 250 patients = 0.040 or 4%

8

Descriptive Statistics

- Measures of Rates and Ratios
  - *Rate: How fast disease occurs in a population.*
  - *Ratio: How much disease compared to standard.*
- Measures of Central Tendency
  - *Central Tendency: How well the data clusters around an average value.*
- Measures of Dispersion (Variability)
  - *Dispersion: How widely your data is spread from the average.*

9

What Makes a Rate?

1. Numerator (top number)
   - *e.g., number of infections*
2. Denominator (bottom number)
   - *e.g., number of patients [proportion]*
   - *e.g., number of patient-days, number of device-days [incidence density/rate]*
3. Time Frame
   - *e.g., day, week, month*

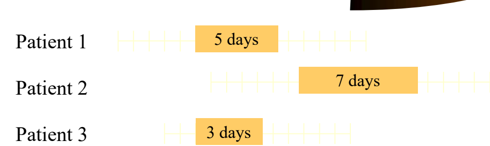
10

Denominators

- Represent the population at risk of becoming part of the numerator
- Often, the most difficult data to obtain, but essential for comparisons
- Ideally, should incorporate time and can account for risk factors such as device use (e.g., device-days), length of stay (e.g., patient-days)

11

What is a Patient/Device-Day?



=15 patient-days, device-days, etc.

- Gives more information than simply—3 patients
- Strategies: e.g., count how many at 9 am

12

Rate Measures

- Prevalence
- Incidence
- Attack Rate

13

Prevalence

- Prevalence: the total number of cases of disease existing in a population at a point in time.
  - *e.g., # of MRSA cases per population* ***on*** *March 8*

$$\frac{\text{Count of existing cases}}{\text{Number of people at risk}} \times \text{constant (e.g., 100 or 1000)} =$$

14

Incidence

- Incidence: the number of new cases of disease in a population over a period of time.
  - *e.g., # of* ***new*** *MRSA cases per population* ***during*** *March*

$$\frac{\text{Count of new cases}}{\text{Number of people at risk}} \times \text{constant (e.g., 100 or 1000)} =$$

15

Attack Rate

- Attack Rate: the number of new cases of disease out of the population at risk.
  - Related to incidence but always uses 100 as the constant, so it is expressed as a percent.
  - Often used for outbreaks or clusters that occur over a short period of time
  - *e.g., % of patients with MRSA during outbreak in Med ICU in March*

$$\frac{\text{Count of new cases}}{\text{Number of people at risk}} \times 100 =$$

16

Example 1:

- You perform HAI surveillance for ventilator associated pneumonias (VAP) and central line associated bloodstream infections (CLABSI) in your 12 bed intensive care unit.
- In March, you identify 2 new VAPs, 4 new CLABSIs and 3 new respiratory infections (not ventilator associated).

17

Example 1:

- The admitting department tells you that in March there were 89 patients in the unit with 311 patient-days.
- Respiratory care tells you that they provided 162 ventilator-days of care to 47 patients in March.
- You count the central line-days and find 284 line-days in 84 patients in March.

18

Example 1:

- In March, what was the VAP rate?
  - Incidence or prevalence?
  - Numerator?
  - Denominator?
  - Units?

19

Example 1:

- In March, what was the CLA-BSI rate?
  - Incidence or prevalence?
  - Numerator?
  - Denominator?
  - Units?

20

Example 1:

- In March, what was overall infection rate?
  - Incidence or prevalence?
  - Numerator?
  - Denominator?
  - Units?

21

Example 1:

- On April 7, you were worried about the BSI rate so you return to the unit to do a "spot check" on all of the patients for a BSI.
- At that time with a census of 12, you reviewed 11 charts and found 1 healthcare associated BSI.

22

Example 1:

- On April 7th, what was the BSI infection rate at the time of your spot check?
  - Incidence or prevalence?
  - Numerator?
  - Denominator?
  - Units?

23

What Makes a Standardized Infection Ratio (SIR)?

1. Numerator (top number)
   =*number of observed infections*
2. Denominator (bottom number)
   =*number of expected or predicted infections*
   - Number of predicted infections = calculated based on your hospital's number of procedures, device days, risk factors, nursing units *compared to a standard infection rate* (e.g., historical data, state data, national data)

24

Predicted Number of Infections

- 2015 as baseline year (2022 for MRSA, Surgery Infections COLO/HYST)
- Logistic regression/negative binomial regression
- Limited patient level risk adjustment
  - facility type, bed size, med school affiliation, types of units.

Details: https://www.cdc.gov/nhsn/pdfs/ps-analysis-resources/nhsn-sir-guide.pdf

SIR Example: CLABSI

$$\text{Predicted CLABSI (\#)} = \frac{\text{NHSN CLABSI rate x central line days}}{1000}$$

Location type	CLABSI (#)	Central line days (#)	NHSN CLABSI rate	Predicted CLABSIs (#)
Medical cardiac	2	380	2.0	0.76
Medical	1	257	2.6	0.67
Med/Surg	3	627	1.5	0.94
Neurosurg	2	712	2.5	1.78
Total	8		------	4.15

$$\text{Overall CLABSI SIR} = \frac{\text{observed}}{\text{predicted}} = \frac{8}{4.15} = 1.93$$

25

Standardized Infection Ratio

- SIR = $\frac{\text{\# observed infections}}{\text{\# predicted infections}}$
- SIR >1.0 → *more infections than predicted*
- SIR <1.0 → *fewer infections than predicted*
- *~LOWER SIRs are BETTER~*

26

SIR Interpretations

- **SIR=1**
- The number of infections is the same as the number of expected infections
- No progress has been made in reducing infections since the baseline period or compared to another standard population (e.g., all NC, all US).

27

SIR Interpretations

- If the **SIR is less than 1**
  - Fewer infections than predicted based on standard or baseline data
  - Infection reduction/prevention compared to standard or baseline data
  - 1 minus the SIR = percent reduction:
    For example, a SIR of 0.80 means that there was a 20 percent reduction from the standard population or baseline time period

28

SIR Interpretations

- If the **SIR is greater than 1**
  - More infections than predicted based on standard or baseline data
  - Infections are increased compared to standard or baseline data
  - SIR minus 1 = percent increase:
    For example, a SIR of 1.25 means that there was a 25 percent increase from the standard population or baseline time period

29

Example 1: SIR

- CLABSI rate = 4 CLABSI/284 line days
- Predicted Infections = 0.57
- What is the SIR?
- How would you explain the SIR to your administrator?

30

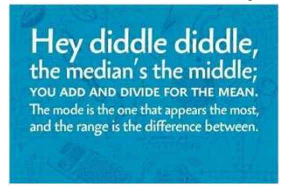
Descriptive Statistics

- Measures of Rates
  - *Rate: How fast disease occurs in a population.*
  - *Ratio: How much disease compared to standard.*
- Measures of Central Tendency
  - *Central Tendency: How well the data clusters around an average value.*
- Measures of Dispersion (Variability)
  - *Dispersion: How widely your data is spread from the average.*

31

Measures of Central Tendency

- **Mean:** average of a group of numbers
- **Median:** middle number in an ordered group of numbers
- **Mode:** most common value in a group of numbers



32

Descriptive Statistics

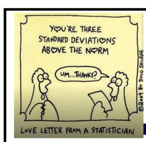
- Measures of Rates
  - *Rate: How fast disease occurs in a population.*
  - *Ratio: How much disease compared to standard.*
- Measures of Central Tendency
  - *Central Tendency: How well the data clusters around an average value.*
- Measures of Dispersion (Variability)
  - *Dispersion: How widely your data is spread from the average.*

33

Measures of Dispersion

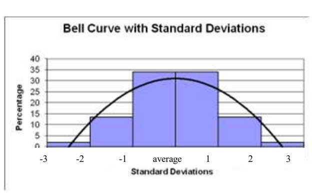
- Range: the largest value minus the smallest value
- Standard deviation: describes the variability or dispersion in the data set

34



Standard Deviation

- **A** measure of degree of variability (spread) in individuals in the sample
  - Standard ("average") deviation ("difference") between an individual's mean and the sample mean
- In a normally distributed data set,



68% of values ± 1 SD

95% of values ± 2 SD

99% of values ± 3 SD

35

Example 2:

- Your administrator is becoming concerned about the impact of healthcare-associated infections on the length of stay in your ICU.
- She has asked you to provide her with some data to confirm her suspicions.

36

Example 2:

- Over the last 3 months you have identified a series of 31 ventilator-associated pneumonias with the total length of stay for each ICU patient as follows:

9, 7, 14, 11, 12, 22, 15, 10, 29, 16, 11, 7, 5, 12, 17, 25, 14, 14, 15, 23, 20, 11, 12, 18, 19, 11, 8, 6, 84, 12, 11

37

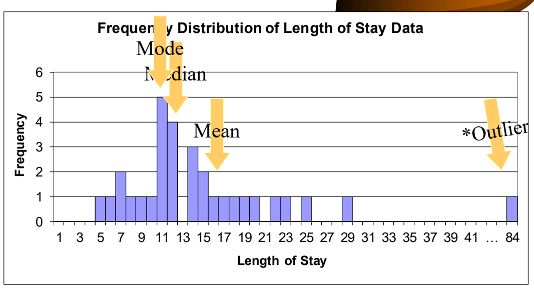
Example 2:

- What is the:
  - Mean?
  - Median?
  - Mode?
  - Range?

HINT: 5, 6, 7, 7, 8, 9, 10, 11, 11, 11, 11, 11, 12, 12, 12, 12, 14, 14, 14, 15, 15, 16, 17, 18, 19, 20, 22, 23, 25, 29, 84

38

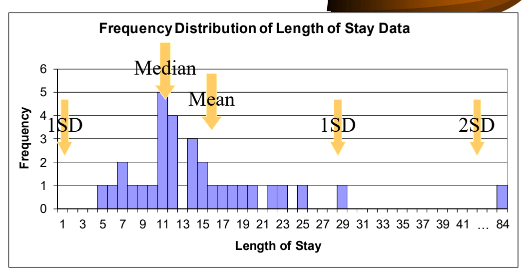
Example 2: Central Tendency



*outlier: a value that falls outside the overall pattern.

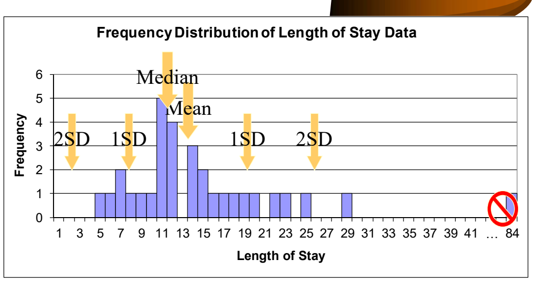
39

Example 2: Dispersion



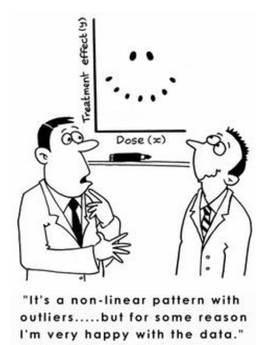
40

Example 2: Dispersion



41

Displaying Surveillance Data



42

Displaying and Interpreting Surveillance Data

- Graphs: a visual representation of data on a coordinate system (e.g., two axes)
- Tables: a set of data arranged in rows and columns

43

Data Types

- Quantitative variables: numerical values
  - *(e.g., number of infections, number of patients)*
- Categorical variables: descriptive groups or categories
  - *(e.g., units in the hospitals, occupational groups)*

44

Features of Graphs and Tables

Graphs and tables should be self-explanatory!

- Clear, concise title: describes person, place, time
- Informative labels: axes, rows, columns
- Appropriate intervals for axes
- Coded and labeled legends or keys
- Use footnotes to:
  - Explain codes, abbreviations, and symbols
  - Note exclusions
  - Note data source

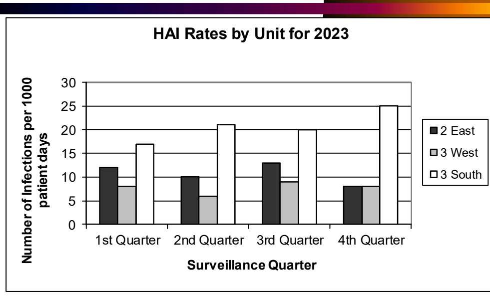
45

Graph Types

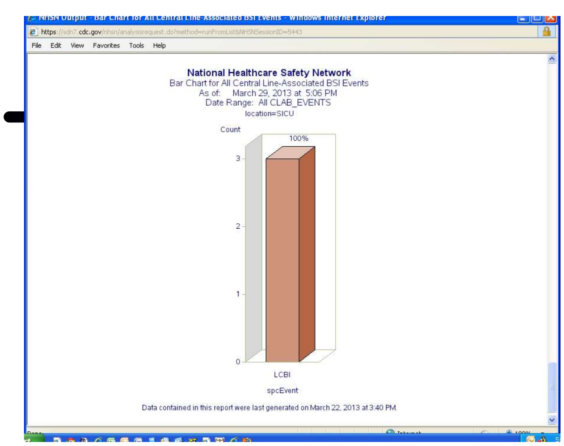
- Bar Graphs
  - *E.g, Histograms (shown in previous example)*
  - *E.g., Comparison between categories*
  - *E.g., Epidemic Curves*
- Line Graphs
  - *E.g., To show trends over time*
- Pie Charts
  - *E.g., As a percentage of a whole*

46

Bar Graph

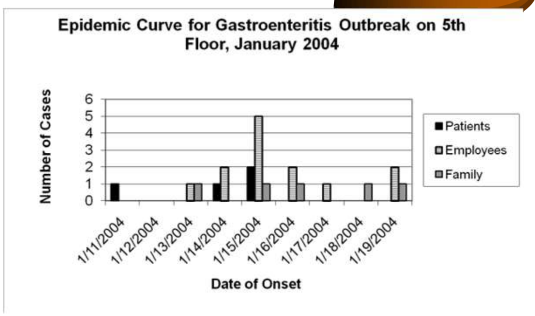


47



48

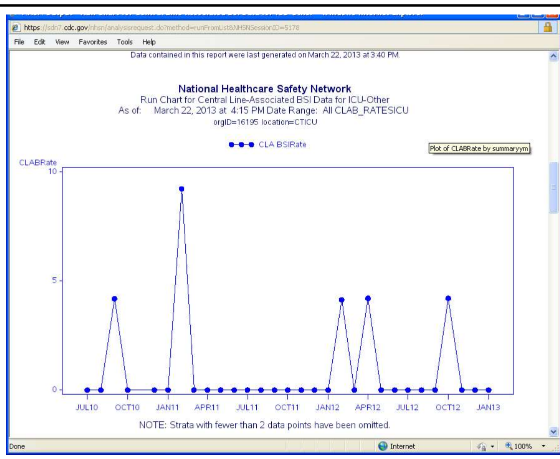
Epi Curve



49

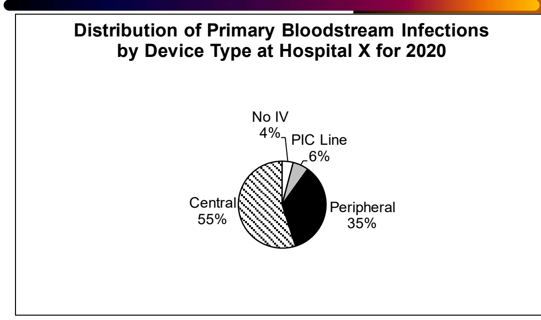
Line Graph

50

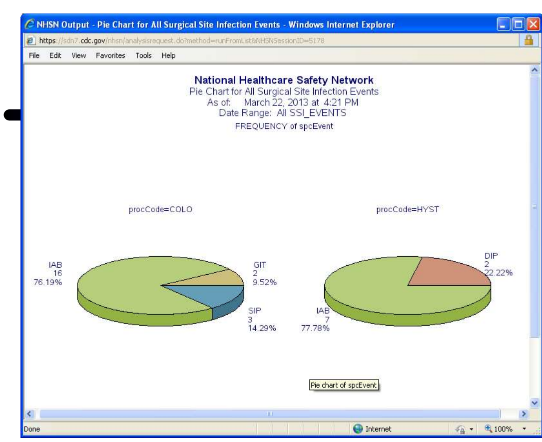


51

Pie Chart



52



53

Tables

Number of Newly Diagnosed Cases by Age, United States, 2021	
Age Group (Years)	**Number of Cases**
0-4	1242
5-14	1081
15-24	2482
25-44	8153
45-64	10916
65+	7124
Total	**30998**

54

NHSN Output - Rate Table for Central Line-Associated BSI Data for ICU-Other - Windows Internet Explorer

https://sdn7.cdc.gov/nhsn/analysisrequest.do?method=runFromList&NHSNSessionID=5178

File Edit View Favorites Tools Help

Source of aggregate data: 2011 NHSN Data
Data contained in this report were last generated on March 22, 2013 at 3:40 PM.

National Healthcare Safety Network
Rate Table for Central Line-Associated BSI Data for ICU-Other
As of: March 22, 2013 at 4:17 PM
Date Range: All CLAB_RATESICU

orgID=16195 loccdc=IN:ACUTE:CC:C

location	summaryYM	CLABCount	numCLDays	CLABRate	CLAB_Mean	IDR_pval	IDR_pctl	numPatDays	LineDU	LineDU_Mean	P_pval	P_pctl
CICU	2010M07	0	223	0.000	1.1	0.7804	25	268	0.832	0.42	0.0000	94
CICU	2010M08	0	290	0.000	1.1	0.7243	25	294	0.986	0.42	0.0000	98
CICU	2010M09	1	236	4.237	1.1	0.2308	96	262	0.901	0.42	0.0000	95
CICU	2010M10	0	276	0.000	1.1	0.7357	25	328	0.841	0.42	0.0000	94
CICU	2010M12	0	253	0.000	1.1	0.7548	25	269	0.941	0.42	0.0000	96
CICU	2011M01	1	282	3.546	1.1	0.2692	93	304	0.928	0.42	0.0000	96
CICU	2011M02	0	298	0.000	1.1	0.7179	25	314	0.949	0.42	0.0000	97
CICU	2011M03	0	241	0.000	1.1	0.7649	25	274	0.880	0.42	0.0000	95
CICU	2011M04	1	238	4.202	1.1	0.2325	95	272	0.875	0.42	0.0000	95
CICU	2011M05	0	213	0.000	1.1	0.7891	25	281	0.758	0.42	0.0000	92
CICU	2011M06	0	237	0.000	1.1	0.7683	25	253	0.937	0.42	0.0000	96
CICU	2011M07	0	161	0.000	1.1	0.8361	25	227	0.709	0.42	0.0000	91
CICU	2011M08	0	218	0.000	1.1	0.7847	25	280	0.779	0.42	0.0000	92
CICU	2011M09	0	195	0.000	1.1	0.8051	25	295	0.661	0.42	0.0000	88
CICU	2011M10	0	239	0.000	1.1	0.7666	25	316	0.756	0.42	0.0000	92
CICU	2011M11	1	230	4.348	1.1	0.2257	96	287	0.801	0.42	0.0000	93
CICU	2011M12	0	228	0.000	1.1	0.7760	25	317	0.719	0.42	0.0000	91

Done | Internet | 100%

55

Interpreting Surveillance Data

56

NHSN data summary, 2013

Urinary catheter associated UTI Rate*					Percentile				
Types of Location: Critical care units	No. Loca-tions	No. of CAUTI	Urinary catheter days	Pooled Mean	10%	25%	50% (median)	75%	90%
Medical cardiac	384	1494	658,345	2.3	0.0	0.7	1.9	3.4	4.9
Medical/surgical ≤15 beds	1645	2429	1,910,118	1.3	0.0	0.0	0.4	1.7	3.1
Surgical cardiothoracic	453	1715	942,852	1.8	0.0	0.7	1.5	2.4	3.4

* Number of CA UTIs x 1000
Number of urinary catheter days

57

What does this NHSN data summary tell you?

- What is the mean UTI rate in the medical cardiac critical care unit?
- If your medical/surgical (≤15 beds) critical care unit has a rate of 1.6 UTIs per 1000 urinary catheter days—between what percentiles is it compared to the NHSN data?
- If your surgical cardiothoracic critical care unit has a rate of 4.2 UTIs per 1000 urinary catheter days—between what percentiles is it compared to the NHSN data?

58

Table 3. State-specific standardized infection ratios (SIRs) and facility-specific SIR summary measures,
NHSN Acute Care Hospitals reporting during 2023
3b. Central line-associated bloodstream infections (CLABSI), critical care locations[1]

State	State NHSN Mandate	No. of Acute Care Hospitals Reporting	No. of Infections: Observed	No. of Infections: Predicted	SIR	95% CI for SIR: Lower	95% CI for SIR: Upper	Facility-specific SIRs: No. of hosp with at least 1 predicted CLABSI	% of hosp with SIR sig higher than national SIR	% o… with … low… nat… S…					
Alabama	Yes	67	184	200.245	0.919	0.793	1.059	30	10%						
Alaska	No	7	12	11.728	1.023	0.554	1.740	3							
Arizona	No	54	129	215.833	0.598	0.501	0.708	37	3%	11%	0.000	0.000	0.363	0.920	1.120
Arkansas	Yes	38	89	118.803	0.749	0.605	0.917	22	5%	5%	0.000	0.338	0.435	0.908	1.543
California	Yes	314	815	973.984	0.837	0.781	0.896	195	10%	4%	0.000	0.266	0.727	1.172	1.899
Colorado	Yes	47	70	111.202	0.629	0.494	0.791	28	0%	4%	0.000	0.000	0.556	0.808	1.297
Connecticut	Yes	28	55	90.588	0.607	0.462	0.784	18	0%	0%					
D.C.	Yes	8	40	55.103	0.726	0.526	0.979	7							
Delaware	Yes	9	18	27.823	0.647	0.395	1.003	5							
Florida	No	204	424	723.701	0.586	0.532	0.644	131	4%	11%	0.000	0.000	0.449	0.807	1.395
Georgia	Yes	91	296	363.551	0.814	0.725	0.911	53	9%	2%	0.000	0.514	0.751	1.074	1.644
Guam	No	2													
Hawaii	Yes	15	34	33.091	1.027	0.723	1.420	9							
Idaho	No	14	14	27.205	0.515	0.293	0.843	7							
Illinois	Yes	118	275	333.760	0.824	0.731	0.926	65	5%	2%	0.000	0.307	0.603	1.137	1.691
Indiana	Yes	70	105	219.359	0.479	0.393	0.577	41	2%	12%	0.000	0.144	0.459	0.774	0.947
Iowa	No	30	54	62.527	0.864	0.655	1.118	10	10%	0%					
Kansas	No	36	39	61.232	0.637	0.459	0.862	12	0%	8%					
Kentucky	Yes	63	148	167.511	0.884	0.750	1.035	29	7%	10%	0.136	0.431	0.928	1.381	1.784
Louisiana	No	70	122	151.943	0.803	0.670	0.955	31	6%	3%	0.000	0.346	0.679	1.055	1.476
Maine	Yes	14	24	25.594	0.938	0.615	1.374	4							
Maryland	Yes	44	124	134.469	0.922	0.770	1.096	32	9%	3%	0.000	0.000	0.693	1.051	1.853
Massachusetts	Yes	57	171	228.248	0.749	0.643	0.868	28	11%	7%	0.000	0.393	0.693	0.961	1.450
Michigan	No	77	258	274.948	0.938	0.829	1.058	52	13%	0%	0.000	0.301	0.676	1.284	1.796
Minnesota	Yes	33	99	142.846	0.693	0.566	0.840	17	6%	12%					
Mississippi	Yes	42	87	87.372	0.996	0.802	1.222	18	6%	0%					
Missouri	Yes	70	159	228.195	0.697	0.595	0.812	32	6%	6%	0.000	0.339	0.603	0.815	1.273
Montana	No	11	5	11.033	0.453	0.166	1.005	5							
Nebraska	No	16	34	43.289	0.785	0.553	1.085	8							
Nevada		20	79	94.974	0.832	0.663	1.031	14	0%	7%					
New Hampshire	Yes	13	12	23.951	0.501	0.271	0.852	7							
New Jersey	Yes	71	103	194.304	0.530	0.435	0.640	50	0%	6%	0.000	0.000	0.445	0.775	1.128
New Mexico	Yes	27	22	42.433	0.518	0.333	0.772	8							
New York	Yes	159	452	565.756	0.799	0.728	0.875	105	8%	4%	0.000	0.000	0.613	1.083	1.798
North Carolina	Yes	83	270	310.729	0.869	0.770	0.977	40	8%	3%	0.000	0.330	0.788	1.253	1.530
North Dakota	No	7	8	14.427	0.555	0.258	1.053	4							

How does NC perform to other states?

HAI Progress Reports | NHSN | CDC
Current HAI Progress Report | HAIs | CDC
Healthcare- and Community-Associated Infections | A.R. & Patient Safety Portal

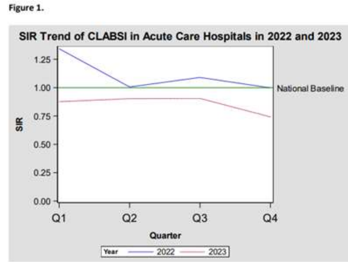
59

NC DPH HAI Reports

D. HEALTHCARE-ASSOCIATED INFECTIONS TRENDS FOR 2022 AND 2023

North Carolina facilities strive to bring the SIR down to below the national baseline and this effort is reflected in the data. See below for how the SIR tracks across the year for 2022 and 2023.

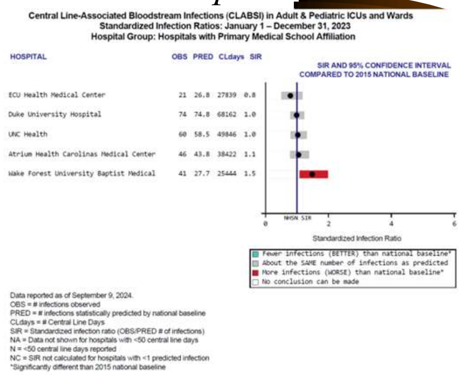
Figure 1.



Interpreting Figure 1:
- The majority of 2022 (Q2, Q3, Q4) and majority of 2023 (Q1, Q2, Q3) experienced about the same number of CLABSIs as predicted, performing the SAME as the 2015 national experience.
- 2022 Q1 experienced more CLABSIs than predicted, performing WORSE than the 2015 national experience.
- 2023 Q4 experienced fewer CLABSIs than predicted, performing BETTER than the 2015 national experience.

60

NC DPH HAI Reports


Central Line-Associated Bloodstream Infections (CLABSI) in Adult & Pediatric ICUs and Wards
Standardized Infection Ratios: January 1 – December 31, 2023
Hospital Group: Hospitals with Primary Medical School Affiliation

HOSPITAL	OBS	PRED	CLdays	SIR
ECU Health Medical Center	21	26.8	27839	0.8
Duke University Hospital	74	74.8	68162	1.0
UNC Health	60	58.5	49846	1.0
Atrium Health Carolinas Medical Center	46	43.8	38422	1.1
Wake Forest University Baptist Medical	41	27.7	25444	1.5

Data reported as of September 9, 2024.
OBS = # infections observed
PRED = # infections statistically predicted by national baseline
CLdays = # Central Line Days
SIR = Standardized infection ratio (OBS/PRED # of infections)
NA = Data not shown for hospitals with <50 central line days
N = <50 central line days reported
NC = SIR not calculated for hospitals with <1 predicted infection
*Significantly different than 2015 national baseline

61

Determine the Significance of Changes to Surveillance Data

62

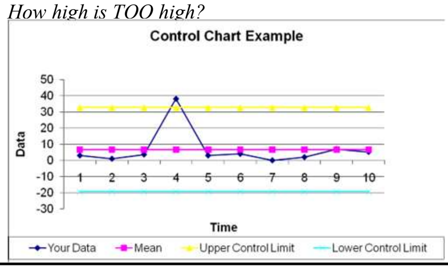
Determine the Significance- How?

- Practical Significance vs. Statistical Significance
- Make comparisons
  - *For example: over time, to other units, to other hospitals (NHSN data)*
  - Remember to choose appropriate data for comparison (*i.e., same denominator units)*
- Apply a type of statistical test
  - *e.g., control charts (for time trends)*
- Other statistical tests and measures
  - P-values
  - 95% confidence intervals

63

Control Charts

- Tool to determine when infection rates are out of range. *How high is TOO high?*



64

Control Chart Example 3:

Month	2015 BSI Rate	Moving Range
1	4.5	--
2	3.2	1.3
3	3.6	0.4
4	3.5	
5	3.0	
6	4.0	
7	4.1	
8	4.6	
9	4.8	
10	5.2	
11	5.7	
12	6.5	

- Find the mean of the BSI rates for the last year.
- Calculate the moving ranges (subtract month 1 from 2, month 2 from 3…) and take absolute values (no negative values).
- Calculate the mean of the moving ranges.

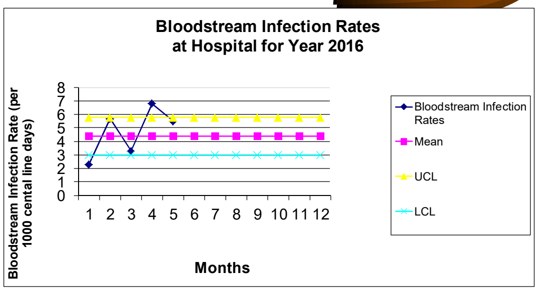
65

Control Chart Example 3:

- Calculate Upper Control limit= Mean + (2.66 x Mean of Moving Range)
- Calculate Lower Control limit= Mean - (2.66 x Mean of Moving Range)
- Draw horizontal lines at the mean, UCL and LCL based on your historical data
- Then graph your current data and use the limits to identify potential problems.

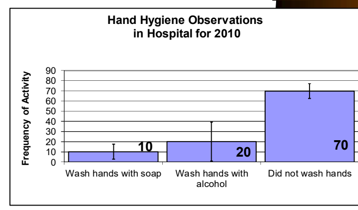
66

Control Chart Example 3:



67

Samples of P values and 95% Confidence Intervals in use



68

Statistical Tests – Why do we need them?

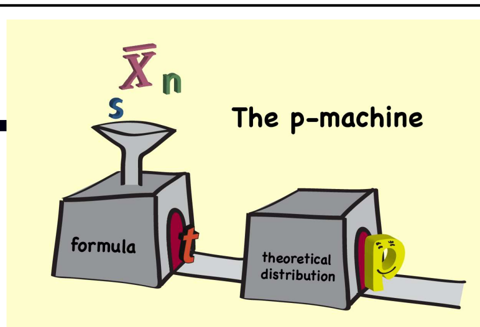
- **Is this real?** Use in clinical research designed to tell if the difference seen is due to chance, or due to some other cause (i.e. a real difference)
- We use these measures to make an **inference**
  - Process of drawing a conclusion about a larger group based on a sample or subset of the group

69

P value

- **P value**: probability of finding a difference as extreme or more extreme than what was found, assuming that the null hypothesis is true
  - Can be used as a measure of the degree of compatibility between observed data and null hypothesis
  - The conventional (yet arbitrary) threshold is 0.05, below which the null hypothesis is rejected
  - 0.05 accepts a 5% risk of a Type 1 error

70



71

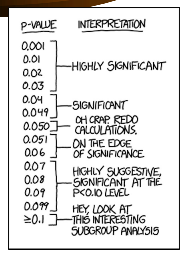
P-Value Example:

- "Our study showed that people who washed their hands were less likely to get sick (P=0.06) and more likely to be nurses (P=0.01)."

72

P-Value Interpretation

- Probability that the difference does not reflect a true difference and is only due to chance.
- e.g., P=0.05 means that 95 out of 100 times your estimate was truly significant
- Generally a level of P<0.05 is considered "statistically significant."



73

Estimation: 95% Confidence Interval

- **95% Confidence Interval (CI):** calculated range of values surrounding the point estimate that are consistent with true effect
  - Formula: point estimate of the mean +/-($2 * s/\sqrt{n}$)
- **Means that you are 95% confident that the true average value lies within this interval.**

74

Statistical Variation of Estimates

- Consider your calculated infection rate to be an estimation of the true rate.

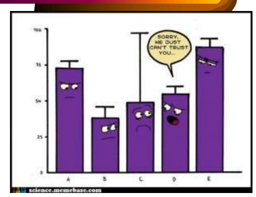
  Why an estimation?

  - You may only do surveillance on a sample of patients in your hospital.
  - If surveillance activities were repeated by other IPs, your numerators may vary slightly based on interpretation of case definitions, available clinical information in the chart, etc.

75

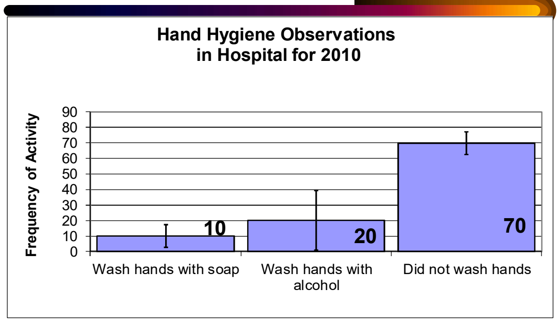
95% Confidence Interval Interpretation

- Confidence interval size:
  - Wide: less confident with that estimate
  - Narrow: more confident with that estimate
- For comparisons,
  - Overlapping intervals suggest no significant difference
  - Non-overlapping intervals suggest significant differences



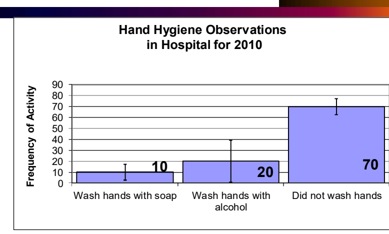
76

95% Confidence Interval Example:



77

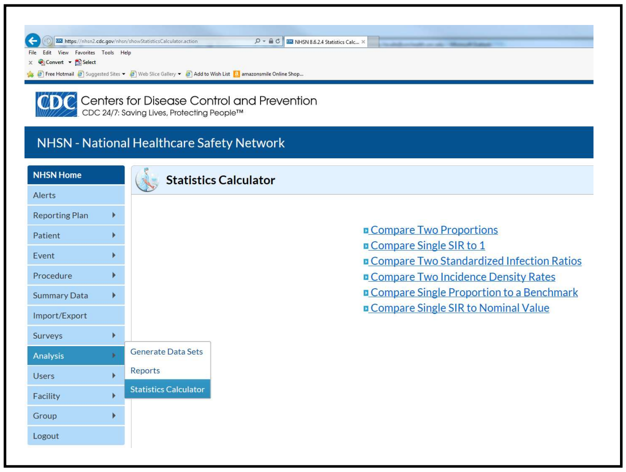
95% Confidence Interval Example:



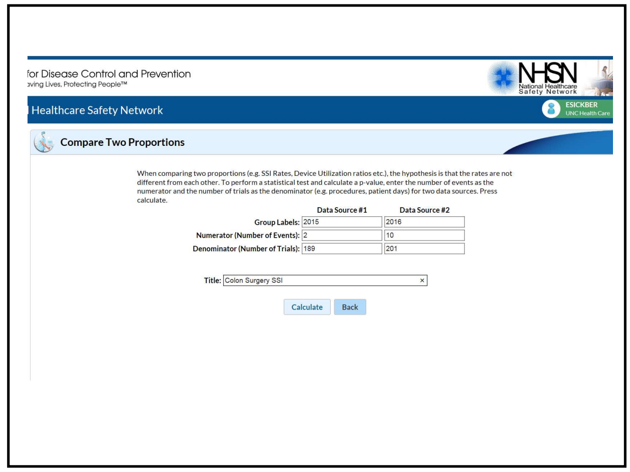
Is the frequency of not washing hands at this hospital statistically significantly different than the frequency of washing hands with soap? *YES – the 95% CI do not overlap*

Is the frequency of washing hands with soap at this hospital statistically significantly different than the frequency of washing hands with alcohol? *NO– the 95% CI overlap*

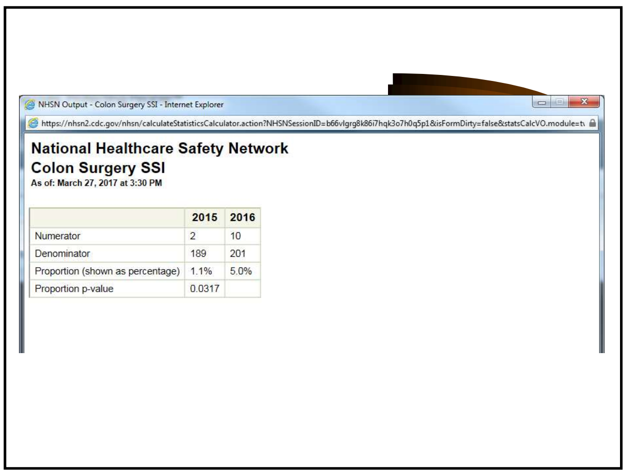
78



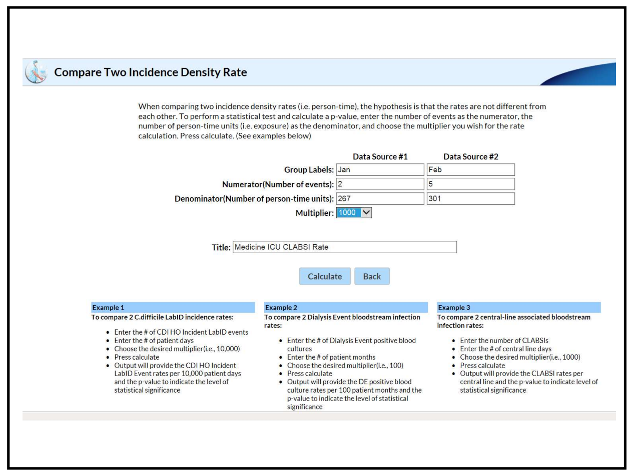
79



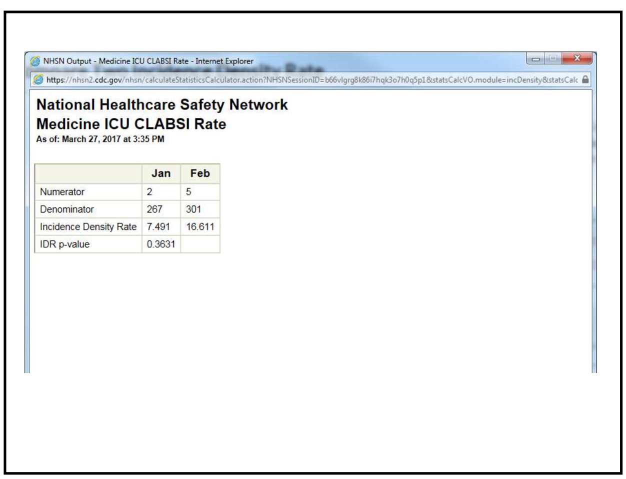
80



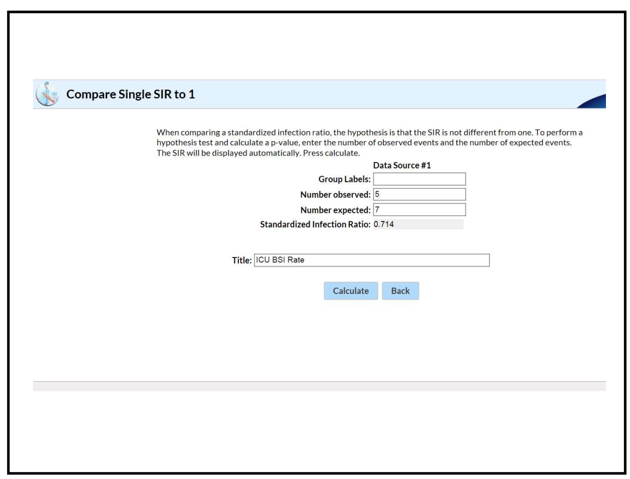
81



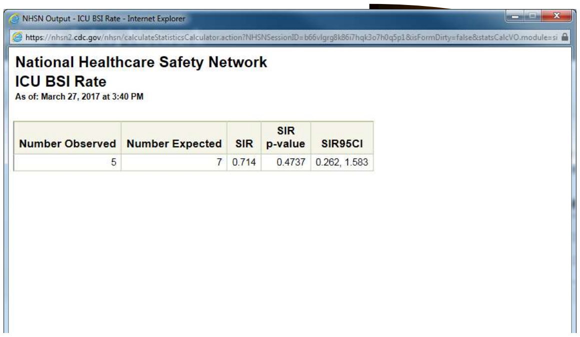
82



83



84



Number Observed	Number Expected	SIR	SIR p-value	SIR95CI
5	7	0.714	0.4737	0.262, 1.583

85

Conclusions

- Describe Surveillance Data
- Display and Interpret Surveillance Data
- Determine the Significance of Changes to Surveillance Data

86

Questions?

87

Group Exercises Using Excel

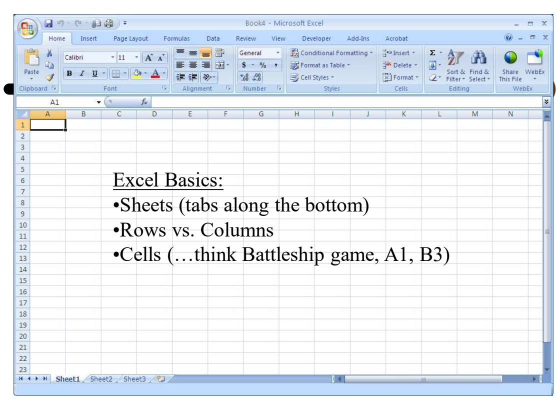
- Infection Rates
  - Create a table
  - Practice formulas
  - Optional activities
    - Graph rates
    - Add 2nd series on graph for NHSN benchmark
    - SIR calculation

88

Group Exercises Using Excel

- Outbreak Investigation
  - Create line-listing of outbreak cases
  - Practice formatting cells, copy/paste, sorting
  - Optional activities:
    - Create a frequency table of cases
    - Graph outbreak epi-curve

89



90

Exercise Wrap-up

- Use Excel as a tool for
  - Calculations of infection rates
  - Creating line-listing for outbreaks or cluster investigations
  - Displaying data graphically
- Use each cell in Excel to capture single piece of data
- Graphs and tables should be self-explanatory!
  - Clear, concise title, informative labels
- Practice, practice, practice!

91

92