Objectives

• Describe two methods to use to analyze surveillance data
• Identify four methods to report data
• Describe how technology supports infection surveillance, prevention and control
How can we use data to improve healthcare practices and outcomes?

First, we must analyze the data....

Data are raw facts and figures: they must be analyzed to create meaningful information that can tell us something.

Then we can use the information...

Meaningful information can guide interventions...
Interventions can reduce infection rates.


UTI Rate per 1,000 catheter-days
How do we use data?

• Tracking and trending
• Benchmarking and comparing
• Improving performance
• Mandatory reporting

Tracking and Trending

Watch for trends & changes over time
• Compare data with previous periods
• Monitor seasonal occurrences
  – RSV, flu, etc.
• Identify outbreaks or clusters of infection
• Detect sentinel event

Benchmarking and Comparing

• Compare our data against others
• Need appropriate benchmark or comparative data
  – Data must be collected and analyzed using same methodology and criteria
• Use to detect areas for improvement
  Ensure that you compare apples to apples
Improving Performance

• Use data/findings to improve performance
  – Promote the best patient/resident outcomes
  – Promote compliance with infection prevention practices
  – Identify risk factors for infection and other adverse events, such as injuries

Measure → feedback data → implement change

Mandatory Reporting

• Used to compare data reported from healthcare facilities

We analyze data using….

• Rates
  – HAIs, compliance with infection prevention practices
• Ratios
  – Device utilization (central line, ventilator, urinary catheter)
  – Standardized infection ratio (SIR)
Comparing Ourselves with Others: The Meaning of our Data and Findings

- The IP is expected to understand how facility rates are reported and compared:
  - How does our organization compare with others?
  - Are groups being compared truly similar?
  - Is the sample size large enough to demonstrate a difference?

Comparing Ourselves with Others: Relevant Statistical Measures

- Some statistical measurements used when comparing groups:
  - Mean (average)
  - Median (midpoint)
  - Range
  - Percentile
  - P-value
  - Confidence interval

Mean, Median and Range

- **MEAN** is the average
- **MEDIAN** is the midpoint or middle value: where 50% of the values are above and 50% are below this number
- **RANGE** is the value that represents the difference between the highest and lowest values in the data set
Mean, Median and Range

Example

Your data set is 1, 3, 6, 10, and 25

• **MEAN** (average) = 9
  - add up all numbers (45), then divide by how many numbers there are (5)
• **MEDIAN** (midpoint) = 6
  - place the numbers in value order and find middle number
• **RANGE** = 24
  - difference between 1 and 25

Percentiles

• Used to indicate relative position of a measurement with respect to other measurements in a set of data
• 50th percentile is the median
  - half of the values in the distribution are lower and half are higher than the median value
• Used in comparative databases (NHSN, state)

Using Mean & Percentiles to Compare Rates

Table 5. Pooled means of the distribution of urinary catheter-associated UTI rates, by type of location, DA module, 2010

<table>
<thead>
<tr>
<th>Type of location</th>
<th>Denominator (catheter-days)</th>
<th>Numerator (URIs)</th>
<th>UTI rate</th>
<th>0.25%</th>
<th>50%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital inpatient</td>
<td>274,236</td>
<td>495</td>
<td>1.83</td>
<td>1.35</td>
<td>2.4</td>
<td>3.5</td>
</tr>
<tr>
<td>Hospital non-inpatient</td>
<td>206,774</td>
<td>366</td>
<td>1.77</td>
<td>1.23</td>
<td>2.3</td>
<td>3.5</td>
</tr>
<tr>
<td>Skilled nursing facility</td>
<td>18,035</td>
<td>30</td>
<td>1.66</td>
<td>1.11</td>
<td>2.1</td>
<td>3.2</td>
</tr>
</tbody>
</table>

* CAUTI rates per 1,000 urinary catheter-days (CAUTI / urinary catheter-days x 1,000)

Using Mean & Percentiles to Compare Device Utilization Ratios

Table 5. Pooled means of the distribution of urinary catheter utilization ratios, by type of location, DA module, 2010

<table>
<thead>
<tr>
<th>Type of location</th>
<th>2009 urinary catheter-days</th>
<th>2009 patient-days</th>
<th>Pooled mean</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Care Units</td>
<td>121,264</td>
<td>47,068</td>
<td>2,050</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Medical Intensive Care</td>
<td>36,022</td>
<td>20,343</td>
<td>1,817</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>Medical Telemetry</td>
<td>16,370</td>
<td>12,757</td>
<td>1,792</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>Medical Surgical</td>
<td>21,250</td>
<td>12,757</td>
<td>1,792</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>Medical Floor</td>
<td>39</td>
<td>203,200</td>
<td>161,300</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Other inpatient units</td>
<td>367 (80)</td>
<td>493,729</td>
<td>405,594</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>

So...how do we use our findings to improve outcomes?

- Examine device-associated infection rates and DU ratios together so that infection prevention measures can be appropriately targeted
  - EXAMPLE: you find that the CAUTI rate for a particular unit is consistently above the 90th percentile and the urinary catheter utilization ratio is routinely between the 75th and 90th percentile.
  - Since the urinary catheter is a significant risk factor for UTI, you should promote practices that:
    - Assess use and remove unnecessary catheters
    - Focus on improving infection prevention strategies

For More Information on Using Percentiles in NHSN Reports to Compare Data see...

  - Appendix A. How to calculate a device-associated infection rate and device utilization ratio with Device-associated Module data
  - Appendix B. Interpretation of percentiles of infection rates or device utilization ratios
How do we statistically compare groups or measurements?

Measures Used to Statistically Compare HAI Rates

• P-value
  – A statistical measure used to determine whether or not there is a significant difference between two rates
• Standardized Infection Ratio (SIR)
  – Used to compare a facility’s risk-adjusted SSI and CLABSI rates with rates in NHSN

P-value: Level of Significance

• We choose level of significance
• 0.05 is traditional cut-off value
  – If p-value is below 0.05 we state that the finding is statistically significant (i.e., there is a difference between two measurements)
  – If p-value is above 0.05 we state that the finding is not statistically significant (i.e., there is no difference between two measurements)
Using SIR and p-value

How do we interpret the SSI SIR?
• During 2009: 524 procedures performed; 13 SSIs identified
  – Based on NHSN 2006-2008 baseline data, 6,687 SSIs were expected/predicted
  – This results in SIR of 1.94 (13/6.687): signifies that our facility identified 94% more SSIs than expected
  – p-value and 95% CI indicate that # observed SSIs is significantly higher than number of expected SSIs


Confidence Interval (CI)

• Used in place of, or in addition to, p-value
• Stated probability usually 95%; can be confident observed value would fall within CI in 95 of 100 studies
• Interpretation:
  – The narrower the interval, the more precise
  – If CI does not include 1.0, then statistical significance implied (p less than 0.05)

Using SIR and Confidence Interval

Table 2, State-specific Standardized Infection Ratio (SIR) for States Using NHSN to Comply With a Legislative Mandate to Report Certain Line-Associated Bloodstream Infections to the State Health Department January 2008 – June 2009

<table>
<thead>
<tr>
<th>State</th>
<th>No. of Simultaneous Reporting</th>
<th>Observed</th>
<th>Predicted SIR</th>
<th>Lower</th>
<th>Upper</th>
<th>95% CI for SIR</th>
<th>Confidence Representation of SIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>30</td>
<td>64</td>
<td>34.25</td>
<td>0.64</td>
<td>0.52</td>
<td>0.87</td>
<td>![Solid Diamond]</td>
</tr>
<tr>
<td>Connect.</td>
<td>30</td>
<td>68</td>
<td>60.46</td>
<td>0.99</td>
<td>0.75</td>
<td>1.30</td>
<td>![Solid Diamond]</td>
</tr>
<tr>
<td>Delware</td>
<td>8</td>
<td>20</td>
<td>33.84</td>
<td>0.88</td>
<td>0.66</td>
<td>0.81</td>
<td>![Solid X]</td>
</tr>
<tr>
<td>Illinois</td>
<td>40</td>
<td>301</td>
<td>333.24</td>
<td>0.80</td>
<td>0.80</td>
<td>1.11</td>
<td>![Open Circle]</td>
</tr>
<tr>
<td>Maryland</td>
<td>48</td>
<td>256</td>
<td>173.85</td>
<td>1.30</td>
<td>1.14</td>
<td>1.49</td>
<td>![Open Diamond]</td>
</tr>
<tr>
<td>Massechus</td>
<td>75</td>
<td>311</td>
<td>311.44</td>
<td>0.58</td>
<td>0.40</td>
<td>0.78</td>
<td>![Open Diamond]</td>
</tr>
<tr>
<td>Nev. Harmp</td>
<td>24</td>
<td>15</td>
<td>22.95</td>
<td>0.87</td>
<td>0.34</td>
<td>0.81</td>
<td>![Open Diamond]</td>
</tr>
</tbody>
</table>

† Solid diamonds=SIR <1.0, solid X=SIR >1.0, open circle=SIR not different than 1.0

How to Present Data

• Use tables, graphs and charts to organize, summarize, and visually display data
• Always provide a written report to interpret your findings and explain their significance

Keys to a Good Presentation

• Decide what results are worth presenting
• Choose the most appropriate format for the information based on:
  – Type of data
  – Your intended message
  – Audience
Never underestimate the importance of appearance!

Tables

- Data arranged in rows and columns
- Show quantitative information
- Display frequency of occurrence

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3. Pooled means of distribution of laboratory-confirmed central line-associated BSI, by type of location, DA module, 2009</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Location</th>
<th>Occurrence (%)</th>
<th>Occurrence (%)</th>
<th>Occurrence (%)</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical Care</td>
<td>59</td>
<td>193</td>
<td>36,305</td>
<td>8.3</td>
</tr>
<tr>
<td>ICU</td>
<td>19,518</td>
<td>194</td>
<td>255,549</td>
<td>9.2</td>
</tr>
<tr>
<td>Med/OCC</td>
<td>35</td>
<td>151,195</td>
<td>35,135</td>
<td>9.7</td>
</tr>
<tr>
<td>Other/Central</td>
<td>290,066</td>
<td>351,208</td>
<td>39,2</td>
<td>7.7</td>
</tr>
</tbody>
</table>

Source: CDC MMWR. Jan 20, 2012

Proper Construction of Table

- Title describes What, Where and When
- Columns labeled
- Rows labeled
- Abbreviations explained

BSI= bloodstream infection; CLABSI=central line-associated BSI; DA= device-associated

Tables are used to organize data for charts and graphs.

<table>
<thead>
<tr>
<th>Month</th>
<th>CLABSI Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>3.4</td>
</tr>
<tr>
<td>Feb</td>
<td>0.0</td>
</tr>
<tr>
<td>Mar</td>
<td>0.0</td>
</tr>
<tr>
<td>Apr</td>
<td>0.0</td>
</tr>
<tr>
<td>May</td>
<td>0.3</td>
</tr>
<tr>
<td>Jun</td>
<td>0.0</td>
</tr>
<tr>
<td>Jul</td>
<td>0.0</td>
</tr>
<tr>
<td>Aug</td>
<td>2.3</td>
</tr>
<tr>
<td>Sep</td>
<td>0.0</td>
</tr>
<tr>
<td>Oct</td>
<td>1.2</td>
</tr>
<tr>
<td>Nov</td>
<td>0.2</td>
</tr>
<tr>
<td>Dec</td>
<td>0.0</td>
</tr>
</tbody>
</table>

ICU CLABSI per 1,000 central line days by month.

Pie Charts

- Display groups of data in proportion to the whole set of data
- Show percentages of the whole

Proper Construction of a Pie Chart

- Has descriptive title
- Displays % of each slice
- Has legend or key to identify each slice
- Has method (color or other) to differentiate each slice
- Shows Total Number (=100%)
- Has explanation of all abbreviations
Sample Pie Chart

Job category distribution of 58 personnel with needlesticks, St. Anne's Hospital, 2011

What does this chart need before it can be distributed?

Bar Charts

- Used to show relationships between different groups or categories
- Can be displayed vertically or horizontally
- Construct chart so it is self explanatory

Proper Construction of a Bar Chart

Job Categories of Workers Reporting Blood/Body Fluid Exposures, Hospital A, 2010 and 2011, Shown as Percent of Total Cases
N=125 in 2010, 113 in 2011

Abbreviations explained
Line Graphs

- Used to show data observed over time
- Three types we use:
  - Simple line graph
  - Run chart
  - Statistical process control chart

Simple Line Graph

- Use to display data observed over time
  - Y-axis: measurements
  - X-axis: time
- Can be used to demonstrate when infection prevention measures were implemented

Run Chart

- Use to track trends and shifts in observed data over time
- Should show a median (midpoint) line
- Need at least 20 data points to reliably detect meaningful variation
- Expect variation—it’s normal!
Statistical Process Control Chart

• Use to statistically measure and plot variation in a set of data
• Strict rules for construction; many types of SPC
  – Mean (average) used for center line
  – Upper and lower control limits plotted
• Need at least 25 data points to statistically detect special cause variation

Goal of Data Presentation

To synthesize and summarize data in a way that is easily understood by the intended audience

Which Graphic Display Shall I Use?

<table>
<thead>
<tr>
<th>Graphic Display</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table</td>
<td>Summarize data: base for constructing graphs and charts</td>
</tr>
<tr>
<td>Line graph</td>
<td>Display &amp; monitor trends in rates or numbers over time</td>
</tr>
<tr>
<td>Bar Chart</td>
<td>Compare size or frequency of different groups</td>
</tr>
<tr>
<td>Pie Chart</td>
<td>Show pieces of a whole</td>
</tr>
</tbody>
</table>

Graphic source: Wikimedia Creative Commons

APIC Text, Chapter 6. Statistical Process Control

Graphic source: Wikimedia Creative Commons
Caution! Avoid Chart Junk

- Keep it simple
- Focus on the data
- Avoid three-dimensional charts and unnecessary elements

Graphic source: Wikimedia Commons

WHICH IS EASIER TO READ?

Good Visual Presentation

- Keep It Simple!
- Tables and graphs should be able to stand alone
- Good tables and graphs will not make up for bad data
- Only create visuals that will enhance your message
- Don’t misrepresent the data!
Components of a Good Surveillance Report
Tailor to recipients and include the following:
• Methods used: who, what, when, where, & why
• Data presentation:
  – Numerator data (raw numbers)
  – Denominator data (# in population)
  – Rates/ ratios, etc., including calculations
  – Use visual displays
• Discussion: explain and interpret data
• Recommendations: actions needed
• Identify stakeholders
• Author(s) and date of report

A Few Words about Technology…
Use computers and information technology to collect, store, sort, retrieve, analyze, display, and report data
USE Technology

- Electronic information available from many sources
  - Have downloaded into IP database
  - Avoid receiving paper reports from lab, admissions, etc.
- Collaborate with IT resources
  - Let them know what you need
  - Influence development of automated reports you need

Use Technology Tools

- Use available computers and systems in your organization
- Consider commercial surveillance programs
  - Data mining & automated surveillance systems
- NHSN web-based program
- Statistical software
  - Commercial & web-based

Surveillance Software Programs

- Advantages
  - Can provide real time data analysis that allows for timely intervention
  - Can support full house surveillance
  - Potential to allow more time for prevention, education, and intervention
- Disadvantages
  - Cost
  - Need IT infrastructure and ongoing support
  - Learning curve to use
Surveillance Software Programs

- Studies have found:
  - computer-based surveillance systems have same or greater sensitivity and specificity as traditional chart review conducted by an IP
- Electronic systems should integrate laboratory, pharmacy, and census data to identify infections and outbreaks
- Technology enhances the role of the IP, it does not replace it.

APIC Resources on Surveillance Technology

- Position Paper: The Importance of Surveillance Technologies in the Prevention HAI
- Starter Questions for Assessing Capabilities of Surveillance Technology & Their Vendors

APIC Resources on Surveillance Technology

Summary

• We must understand how to analyze, report, and use surveillance data
• Surveillance data should be used to improve healthcare practices and outcomes
• We should optimize the use of computers and information technology to support infection surveillance, prevention and control activities

Resources/Bibliography

• APIC Text of Infection Control and Epidemiology, 3rd ed.
  – Chapter 3. Surveillance
  – Chapter 5. Use of Statistics
  – Chapter 6. Statistical Process Control
  – Chapter 7. Risk-Adjusted Comparisons
  html
• NHSN Home Page: www.cdc.gov/nhsn

Any Questions?