10 Most Common Mistakes IP Make with Data - and How to Avoid Them

Irena L. Kenneley, PhD, APRN-BC, CIC
Katherine Ward, RN, BSN, MPH, CIC
Marilyn Hanchett, RN, MA, CPHQ, CIC

Learning Objectives

After this program, participants will be able to:
• Integrate data management into an effective infection program using a systems, rather than task specific, approach

• Identify common obstacles to using data and offer practical solutions

• Support skill building and professional development of proficient IPs and support their transition into the advanced stage of practice
Infection Preventionists Often Report Feeling Overwhelmed by Data

Why?
• The need to collect
• The need to validate
• The need to understand and prioritize
• The need to share
• The need to publicly report
• The need to compare

Collecting and Using Data is Part of a Larger Process

Data → Information → Knowledge
Improving Your Use of Data Requires BOTH Skills and Knowledge

Skills
- Measures & metrics
- Basic Statistics
- Data Display Tools

Knowledge
- Potential Bias
- Limitations
- Interpretation

Optimum Data Use

Data Use: a Core Competency for Infection Preventionists

Visit the APIC web site for more information on the competency model
Correct Data Use: Essential for Implementing Science

Basic or Empirical Evidence (Non-Experimental Designs)

1. Case studies and other descriptive reports
   - descriptive, exploratory or explanatory analysis of a person, group or event
   - An explanatory case study is used to explore causation in order to find underlying principles

2. Consensus statements/Expert opinion
   - Consensus methods are being used increasingly to solve problems in medicine and health
   - Their main purpose is to define levels of expert agreement on controversial subjects

3. Cross-sectional and prevalence studies
   - involves data collection from a population, or a representative subset, at one specific point in time
Evidence Identified via Studies
(Experimental Designs)

1. **Randomized Control Trials (RCT)**
   - specific type of scientific experiment, and the gold standard for a clinical trial
   - often used to test the efficacy or effectiveness of various types of medical intervention within a patient population

2. **Cohort Studies**
   - often undertaken to obtain evidence to try to refute the existence of a suspected association between cause and effect
   - failure to refute a hypothesis often strengthens confidence in it

3. **Case-Controlled Studies**
   - often used to identify factors that may contribute to a medical condition by comparing subjects who have that condition/disease (the “cases”) with patients who do not have the condition/disease but are otherwise similar (the “controls”)

---

**FAQ:** If RCTs are so important, why are there so few supporting infection prevention practices?
<table>
<thead>
<tr>
<th>Study Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Sectional Prevalence</td>
<td>Confounding factors uncontrollable</td>
<td>Linking exposure with unexposed persons</td>
</tr>
<tr>
<td></td>
<td>Simple Analysis</td>
<td>Average exposure masks complicated events of the disease</td>
</tr>
<tr>
<td></td>
<td>Easy presentation</td>
<td>Population level effect is not associated with individual level effect</td>
</tr>
<tr>
<td></td>
<td>Hypotheses friendly</td>
<td>Can be conducted rapidly and with minimal resources</td>
</tr>
<tr>
<td></td>
<td>Can be conducted rapidly and with minimal resources</td>
<td>Population level effect is not associated with individual level effect</td>
</tr>
<tr>
<td></td>
<td>Internal comparisons of exposure and population status</td>
<td>Must start with clear definition of event</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Drop out rates affect validity</td>
</tr>
<tr>
<td>Cohort</td>
<td>Multiple data collect tools</td>
<td>Can be expensive</td>
</tr>
<tr>
<td></td>
<td>Estimation of disease risk</td>
<td>Time consuming</td>
</tr>
<tr>
<td></td>
<td>Marginalize exposure and selection bias</td>
<td>Complicated and difficult to carry out</td>
</tr>
<tr>
<td></td>
<td>Can study rare and multiple effects</td>
<td>Exposures can be misclassified</td>
</tr>
<tr>
<td></td>
<td>Temporal relationship between factor and outcome is known</td>
<td>Follow-up can be difficult to avoid drops outs</td>
</tr>
<tr>
<td>Case-Control</td>
<td>Quick and less expensive</td>
<td>Built in bias is common</td>
</tr>
<tr>
<td></td>
<td>Studies rare events</td>
<td>Control group difficult to select</td>
</tr>
<tr>
<td></td>
<td>Can study multiple risks and exposures</td>
<td>Can provide ratios but unable to calculate rates</td>
</tr>
<tr>
<td></td>
<td>Study is not time sensitive</td>
<td></td>
</tr>
</tbody>
</table>

## Publications Differ in How Data are Used and Presented

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Scholarly Journal or Peer Reviewed or Refereed</th>
<th>Technical/Trade Journal</th>
<th>Popular Magazine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of Sources</td>
<td>Sources and references are always cited in scholarly articles—with footnotes or a bibliography.</td>
<td>Sometimes sources and bibliographies are given. This varies depending on the publication.</td>
<td>Articles rarely, if ever, cite sources in a bibliography.</td>
</tr>
<tr>
<td>Overall Purpose</td>
<td>To make the information available to the rest of the scholarly world.</td>
<td>Report on trends in the profession and give practical advice to professionals and other interested readers.</td>
<td>To entertain, provide news, provide information about a hobby or special interest (e.g., gardening, motorcycles) promote a viewpoint, or sell a product.</td>
</tr>
<tr>
<td>Review Process</td>
<td>Scholarly articles go through a peer review (referee) process where other scholars in the field evaluate the content of the article. The articles are edited for grammar, format, etc.</td>
<td>Articles are reviewed by the magazine's editorial staff including copy editing for spelling, grammar, etc. Some trade journal articles go through some sort of peer review process—see the publication's website (or editorial page) to find out.</td>
<td>Articles are reviewed by the magazine's editorial staff, including copy editing for spelling, grammar, etc.</td>
</tr>
</tbody>
</table>

RPIC ANYWHERE
ONLINE EDUCATION CENTER
Example: Using Data and Information from APIC Publications

How do you use data from AJIC . . .

Compared to Prevention Strategist?

How Do You Classify . . .

This HICPAC Guideline contains a great deal of data and information. It is an example of (pick one):
Quick Tips for Research Articles

• Assumption of applicability of results
• Conflict of Interest statement
• Use of percentages
  – Without actual numbers

Example:
A study reports that 50% of healthcare workers performed hand hygiene

But was that 2 out of 4 or 200 out of 400?

Quick Tips for Research Articles

• Result section
  – Only significant results
• Exclusion of data with negative results
• Reporting only p-value<0.05
  – Actual p-value result

FAQ: What should I do if I don’t understand how the results/data are described?
Detective Tools

• Be skeptical
• How does this Study apply to my population
• What is not known?
• Was there bias?
• How long was the Study?
• Were the results sustainable?

FAQ: are studies sponsored by industry inherently biased? How can I tell?
Getting the Most from Data Display Tools

• A visual representation or “picture” of the data
• Many types available, including levels of complexity
• Graphs generally have higher visual impact than tables
• Important to use the display method that best describes the point you make with your data
• Use of color not always necessary for high impact

Common Errors with Data Display Tools

• Decimal points
  – Never more than needed
• Lack of concise message
• Express versus Impress
Displaying Data: Tables versus Graphs

Tables
- Individual values
- Precise values
- Multiple units of Measure

Graphs
- Convey information by shape of data
- Ability to show relationships among many values

Common Errors with Graphs

- Use of unnecessary dimensions
  - Strong grid lines
  - Long slanted labels
  - Non-data elements

Too much information is just as bad as not enough!
Common Errors with Graphs

• Use of scales
• Equal spaced marks for uneven intervals
• Too many tick marks and labels
  – Data may be lost
• Too few tick marks and labels
  – Data may be hard to locate
• Not using zero baseline – bar charts

What is Wrong with These Graphs?

No Vertical Axis

No labels on axis
Uneven scale

Vertical axis not at zero
Constructing Tables

- Concise titles
  - Describe content
- Comprehension without referring to text
- Numeric assigned based on order in text
- Abbreviations
  - Footnote for explanation
- Units of measure
  - Column or Row

How Often Have You Seen a Table Like This?

*Does it help you understand the data?*

**ASK:**
- Who is the audience?
- What is the purpose of the table?
- How much detail is needed to achieve the purpose?
- Where will the table be used/how much time do I have to present it?

Age-Adjusted Death Rates
Basic Statistical Tools
Every IP Should Be Able to Use

The P Value (in plain English)
Statistic relating whether or not the sample supports the tested hypothesis

What You Need to Know:
• P values are most often set at .05
• Results that are less than .05 are deemed statistically significant because . . .
• .05 = 95% confidence that the result did not occur by chance alone

What if the P Value is Not .05?

FAQ: the study reports using a P value of .005 rather than .05. Why is this important?

• The smaller the P Value, the less likely the results obtained were due to chance.
• .005 means there is less than 1% chance the results are due to chance
• The use of a P Value less than .05 is at the discretion of the researcher
• .05 is not necessarily better/worse than .005 – it depends on what is being studied
The P Value is Often Used in Infection Prevention and Control

The hospital has changed its patient room cleaning and disinfecting protocol due to an increase in C difficile infections. The Infection Prevention team has completed a study to help determine if the new procedure is beneficial. This is especially urgent because the new procedure costs more and requires more time by EVS staff.

The team’s analysis is ready to be shared with the hospital’s Infection Prevention and Control Committee. The P value for their study was .05. Their result is \( p = 1.273 \).

**What should the infection prevention team report to the committee?**

---

**Another Statistical Tool: Using Averages**

**Average or the Mean**

*Classic bell shaped curve AKA a normal distribution*

- The mean is based on a normal distribution of your data
  - *But ask*
  - Do my data really look like this? If not, how reliable is an average number?
Helpful Tip: Include the Standard Deviation (SD) when You Work with Mean Scores

The SD helps you do two important things with your data:
1. Understand how much variation there is from the mean in your data set (may be called dispersion). A means score used alone may “hide” this variation.
2. Helps measure the confidence in the conclusions you draw from your statistics.

So . . . How Much Data Variation is Acceptable?

General Rule:
The smaller your SD, the less variation in your data
AND
The more your data look like a normal distribution.
Try This Example . . .

The IP has completed a series of education programs and summarized both the mean and standard deviation for 4 groups of participants.

Which set of scores indicates the most consistent level of performance among the attendees?

What is Correlation?

Correlation describes a statistical relationship between two sets of data. Reported between -1 through +1.

- 0 means there is no correlation at all
- The closer the result is to +1, the stronger the positive correlation
Try This Example . . .

The IP is analyzing the CAUTI rates reported by two inpatient units/wards of the hospital. Unit 1 thinks its CAUTI rate is significantly different than Unit 2.

The IP is using 3 months of data from both units and has obtained a correlation result of .79

How should the IP interpret and report this result?

Putting Data to Work: Clinical Practice and Decisions

The IP has reviewed two studies reporting the results of a new vaccine. In the first study, 2500 adults in the United States ages 21 – 35 showed 99% protection against the target disease.

In the second study, 75% of adults in US nursing homes showed 88% protection against the target disease.

What conclusion should the IP make from this information?
The Pros and Cons of Data Extrapolation

**PRO**
- Useful when studying individuals who are similar to the original study group
- Helpful in understanding the aggregate effect on a large group/population
- Can be quick, fairly easy and inexpensive

**CON**
- Misleading when used with different populations
- Difficult to try to extend results to situations where actual data are lacking (easy to make a mistake!)
- Assumes past trends will continue, not sensitive to changing situations

Use your Detective Skills!
How Would Respond to This Scenario?

Data: Needlestick injuries have increased in the hospital by 22% in the past 10 months.

Conclusion: This is a serious employee health threat that must be addressed by the IP and Safety Officer immediately. House-wide inservice training must begin in 24 hours.
Here is Another . . . . What Would You Do?

The nurse manager is considering replacing Product A with Product B, which is newer, has advanced design features and costs more. She has read one published article about Product B that favorably described a before/after analysis conducted at another facility. In addition, the manager has heard excellent reports from colleagues at two other hospitals who switched to Product B. The manager now recommends replacing Product A, saying that the data support this decision.

Do you agree? Why or why not?

How Do I Improve My Data Skills?

• Read and critically analyze journal articles. Practice, practice, practice!
• Brush up on basic statistics (not as scary as you may think)
• Experiment with different data display techniques (there are many); get comfortable using software to do this
• Attend ed programs sponsored by APIC Chapters and at the annual conference
• Learn more from APIC webinars
Questions