RESEARCH 101 TRAINING
How to be Smart About Research in 8 Hours or Less

Session 3: Data analysis, interpretation, and presentation

July 13, 2017 @ 5:30-7:00 p.m. Eastern
The 3rd of 3 Sessions

1. Intro and research design
2. Research instruments and data collection
3. Data analysis, interpretation, and presentation
Agreements

- Take care of yourself & ask for what you need
- Ask questions (some research lingo may be unfamiliar and we are here to break it down)
- Step up, step back (draw out others’ ideas)
- Please try to stay off cell phone and social media
- Acknowledge everyone’s experience is different
- “One mic” (please mute phone when not speaking)
- Session is being recorded
Today’s Agenda

1. Analysis and Presentation
2. Interpretation and Conclusions
1. Analysis and Presentation
Data Analysis – the Patients’ Role

• Unlikely you will be running the analysis entirely by yourself!

• But, you may have the opportunity to:
  • Code unstructured data
  • Have input into the types of analysis to run
  • Review the analysis results, and help interpret the findings
  • Suggest changes to the way the analysis is run

• To do those things, you need a basic understanding of how to read and interpret tables, charts, and other types of data analysis output
Research Steps

- Research Question
- Method Selection & Sample Selection
- Data Collection
- Data Analysis & Presentation

Research Hypothesis

Informed best guess about what a study may find

- Level of education influences smoking behavior
- Weather has an impact on joint pain

You can have multiple hypotheses

- Medicine A has better outcomes than medicine B for chronic pain
- Medicine A has fewer side effects

Test, Test, Test

- Design your research study to test whether your hypothesis is true. Do your observations support your hypothesis or not?
- Test your hypothesis using statistical methods
- Draw conclusions
Quantitative Data

• Examples: Rates (incidence, prevalence); close-ended questions
• Discrete Data: data in whole numbers (e.g., 47 years old)
• Continuous Data: measures on a continuous scale (e.g., 47.32 years old)
• Categorical: Descriptive data (no ordering, categories, preferences)
• Ordinal: data that has some kind of ordering like age or income level
Qualitative Data

Qualitative data is based on narrative information, not numerically ‘measurable’ information (e.g., “What does age 47 feel like?”)

- Perceptions
- Quotations
- Experience/Observations
- Open ended questions
Health researchers want to understand the effect of a medication on people with a certain condition

**Quantitative methods**
1. Complete a survey of their symptoms before and after taking the medicine.
2. Measure their symptoms before and after taking the medicine.

**Qualitative methods**
1. Interview patients & ask them to talk about their experience when taking the medicine.
2. Conduct social media discussion about patient experiences

Tools to Collect and Present Data

- Data can be presented in different formats that include tables, graphs, narratives.

- Statistics is a tool that is used for quantitative research.

- Qualitative research uses non-statistical tools.

- Graphs can be used to present both qualitative and quantitative data.
Graphs: A Useful Tool to Communicate

With the Advent of HAART, More People Are Living with HIV Infection (red) as Rates of AIDS-Related Deaths Decline (blue)

Data File Formats
Quant Data File Example

- Columns represent questions (variables)
- First row shows question (variable) names
- Remaining rows represent individual “cases” (participants, respondents)

<table>
<thead>
<tr>
<th>ID</th>
<th>Drugprice</th>
<th>Drugtype</th>
<th>RA</th>
<th>OA</th>
<th>Whypharm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$99</td>
<td>Generic</td>
<td>Yes</td>
<td>No</td>
<td>I like the convenience</td>
</tr>
<tr>
<td>2</td>
<td>$99</td>
<td>Generic</td>
<td>Yes</td>
<td>Yes</td>
<td>Cheap, easy to get there</td>
</tr>
<tr>
<td>3</td>
<td>$99</td>
<td>Generic</td>
<td>No</td>
<td>Yes</td>
<td>Convenient, friendly</td>
</tr>
<tr>
<td>4</td>
<td>$99</td>
<td>Generic</td>
<td>No</td>
<td>Yes</td>
<td>Open 24 hours</td>
</tr>
<tr>
<td>5</td>
<td>$199</td>
<td>Brand</td>
<td>Yes</td>
<td>No</td>
<td>Near my house</td>
</tr>
<tr>
<td>6</td>
<td>$97</td>
<td>Generic</td>
<td>No</td>
<td>No</td>
<td>Free delivery</td>
</tr>
</tbody>
</table>
Data File Formats

Qualitative

- Transcripts (from focus group discussion or interviews)
- Audio and video recordings
- Social media posts

Social Media Chat: Can Pets Help People with Chronic Illness?

Getting my puppy has made so many days so much better.

My dogs are my heating pads, my laughter, my secret keepers, my exercise, my bed warmers, my heart and soul. They're there at night when I need to cry over the pain without judgment.

My cats, horse, and pup give me a reason to get out of bed on the tough days.
Qual Data Analysis: What is Coding?

- Not computer programming “coding”

- Rather: Sifting through verbatim responses in patients’ own words, and “bucketizing” the results according to themes you detect

- Used for both qualitative and quantitative research

- Patient insight can be very useful here
  - But also could introduce bias, so be aware
## Coding - Example

<table>
<thead>
<tr>
<th>ID</th>
<th>Whypharm</th>
<th>Convenient</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I like the convenience</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cheap, easy to get there</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Convenient, friendly</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Open 24 hours</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Near my house</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Free delivery</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Low prices</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Can walk there</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Coupons</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>In my grocery store</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
What do you think of this website?

- The graphics are appealing
- The material is interesting
- I couldn’t find what I was looking for
- The font size is a bit too small
- I like the variety of information
- The information is presented in a way that is easy to understand
- I like the overall look
- It’s hard to see on my smartphone
Qualitative Analysis - Word Clouds

Source: BJC Health Connected Care
Basic Analysis Methods
Data Analysis

- To analyze quantitative data we use **statistics**
- Research question(s) and data types dictate analysis and statistical methods

\[
\begin{align*}
N &= \text{Population size} \\
n &= \text{Sample size} \\
p &= \text{Statistical significance} \\
(p &= \text{How likely differences between groups are due to chance})
\end{align*}
\]

- \( p < 0.05 \) means less than a 5% likelihood (probability) that study findings are due to chance (and therefore, a greater than 95% likelihood that the interventions are related to the results)
Statistical Significance – A Couple of Caveats

- **Statistical significance can be manipulated:**

- **Statistically significant differences are not always important research findings**
  - *Clinical significance* tells whether research findings are of practical use to patients and clinicians
Measures of Central Tendency

- Best for understanding “typical”, but still need to understand distribution/range
  - Use t-tests to compare results across groups

Example: 2, 2, 2, 4, 6, 8, 9

<table>
<thead>
<tr>
<th>Mean</th>
<th>The arithmetic average of a series of values</th>
<th>4.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>The middle value in a series</td>
<td>4</td>
</tr>
<tr>
<td>Mode</td>
<td>The value that occurs most often in a series</td>
<td>2</td>
</tr>
</tbody>
</table>
A way of summarizing data by describing how spread out the data is.

The average score in a class on a math test may be 65 points. But we may also want to talk about the higher scores (above 65) and the lower scores (below 65).

Measures of Spread can be calculated using different statistical methods such as Standard Deviation and variance.
Data: Tables & Graphs

Frequency Data Tables
The number of observations in each category is called frequency.

<table>
<thead>
<tr>
<th>Weight in grams</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-19</td>
<td>3</td>
<td>10%</td>
</tr>
<tr>
<td>20-39</td>
<td>14</td>
<td>15%</td>
</tr>
<tr>
<td>40-49</td>
<td>8</td>
<td>30%</td>
</tr>
<tr>
<td>50+</td>
<td>12</td>
<td>45%</td>
</tr>
</tbody>
</table>

Graphs
X-Axis = Categories
Y-Axis = Frequency
Distribution Example & Discussion

How would you describe the results of this study?

<table>
<thead>
<tr>
<th>Drug Price ($)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; $100</td>
<td>80</td>
</tr>
<tr>
<td>$100-$199</td>
<td>260</td>
</tr>
<tr>
<td>$200+</td>
<td>80</td>
</tr>
</tbody>
</table>

Price Paid for a One-Month Prescription (n=420)
Central Tendency Example & Discussion

Now, how would you describe the results of this study?

Price Paid for a One-Month Prescription
(n=420)

Mean = 153
How would you describe the results of this study?

Change in Perception of Pain After 3 Months

<table>
<thead>
<tr>
<th></th>
<th>Exercise Group (n=300)</th>
<th>Control Group (n=300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>No change</td>
<td>14%</td>
<td>66%</td>
</tr>
<tr>
<td>Decreased</td>
<td>76%</td>
<td>22%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
2. Interpretation and Conclusions
Describing the data is the first step

Interpretation involves relating results to the research questions, and sometimes, to previous research

Ultimately, what do you recommend based on the results?

- How might this affect the target population? (Patient perspective can be key here)
- What other research is needed?
There are a lot of things we didn’t cover, or just touched on

- PCORI 101 Training Booklet covers a lot of good material – some overlap with this training, and some new
- We plan to continue the conversation…
Thank You!
Useful Resources


- Gapminder.org: This resource has some very creative graphs and resources to use statistics: [http://www.gapminder.org/](http://www.gapminder.org/)

- CDC: Data and statistics [http://www.cdc.gov/datastatistics/](http://www.cdc.gov/datastatistics/)


- Qualitative Health Research (QHR): This is a peer reviewed journal on qualitative research: [http://qhr.sagepub.com/](http://qhr.sagepub.com/)
‘Extra Credit’

Graphics inspiration