MDICx webinar series

From Stories to Evidence: Quantitative patient-preference information to inform product-development and regulatory reviews

Shelby Reed
Professor, Duke School of Medicine

F. Reed Johnson
Professor, Duke School of Medicine

Juan Marcos Gonzalez
Assistant Professor, Duke School of Medicine

March 15, 2018
From Stories to Evidence: Quantitative patient-preference information to inform product-development and regulatory reviews

MDIC Webinar II
March 15, 2018

F. Reed Johnson
Professor, Duke School of Medicine

Shelby Reed
Professor, Duke School of Medicine

Juan Marcos Gonzalez
Assistant Professor, Duke School of Medicine
“Qualitative or quantitative statements of the relative desirability or acceptability of attributes that differ among alternative interventions.”
Preference-Elicitation Approaches

• Qualitative methods (focus groups, public meetings)
  – Identify areas of concern
  – Provide context for product-development and regulatory decisions

• Simple quantitative methods (ranking, threshold)
  – Prioritization
  – Tradeoffs involving only two outcomes

• More advanced quantitative methods (choice experiments, best-worst scaling)
  – Tradeoffs involving more than two outcomes
  – Statistical preference measures (risk tolerance, minimum acceptable benefit, time equivalents)
  – Publishable regulatory-quality evidence

• Today’s focus: discrete-choice experiments
Choice-Experiment Features

- Also known as choice-based conjoint analysis
- Alternatives consist of combinations of features
- Preferences among alternatives depend on the relative importance of features
- Respondents indicate choices among hypothetical alternatives
- Statistical analysis of pattern of choices indicates relative importance of features
Example Choice Question: Parkinson’s

<table>
<thead>
<tr>
<th>Treatment Feature</th>
<th>Medicine Pump</th>
<th>Brain Stimulator</th>
<th>Oral Medicine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect on control of movement symptoms</td>
<td>8 hrs sleeping</td>
<td>8 hrs sleeping</td>
<td>8 hrs sleeping</td>
</tr>
<tr>
<td></td>
<td>4 hrs treatment working</td>
<td>3 hrs treatment working</td>
<td>6 hrs medicine working</td>
</tr>
<tr>
<td></td>
<td>12 hrs treatment working</td>
<td>13 hrs treatment working</td>
<td>10 hrs medicine working</td>
</tr>
<tr>
<td>Ability to find the right words when speaking</td>
<td>No difficulty</td>
<td>A lot of difficulty</td>
<td>Some difficulty</td>
</tr>
<tr>
<td>Ability to think clearly</td>
<td>No difficulty</td>
<td>A lot of difficulty</td>
<td>Some difficulty</td>
</tr>
<tr>
<td>Additional oral medicine you take</td>
<td>None</td>
<td>3 pills 4 times a day</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>(12 pills per day)</td>
<td>(24 pills per day)</td>
<td></td>
</tr>
<tr>
<td>Which would you choose?</td>
<td>Medicine Pump</td>
<td>Brain Stimulator</td>
<td>Stay On Oral Medicine</td>
</tr>
</tbody>
</table>
Discrete-Choice Experiments to Quantify Patient Preferences

• Developed, tested, and validated over past 40 years in
  – market research
  – transportation planning
  – environmental economics
  – health

• Daniel McFadden received the Nobel Prize in Economics in 2000 for conceptual and statistical foundations

• Increased interest and regulatory support because of commitment to patient-centered healthcare
“FDA understands that patients and care-partners who live with a disease or condition ... may have developed their own insights into and perspectives on the benefits and risks of devices reviewed.”

- Voluntary submission of patient-preference data
- Recommendations for collecting patient-preference data for FDA reviews
- Recommendations for including patient-preference information in labeling
Other Sources of Information

- ISPOR Conjoint Analysis Task Force Reports (published in *Value in Health*)
  - Checklist
    - https://www.ispor.org/workpaper/ConjointAnalysisGRP.asp
  - Experimental designs
  - Analysis

- MDIC Framework
  - Report
1. Research question

2. Attributes and levels

3. Construction of tasks

4. Preference elicitation

5. Instrument design

6. Experimental design

7. Data collection

8. Statistical analyses

9. Results and conclusions

10. Study presentation

ISPOR Checklist for Stated-Preference Applications in Medicine
Research Question Considerations

• Study perspective
  *Who? What? Why?*

• Decision-making context
  *Is the decision preference-sensitive?*

• Tractability
  *Can the question be answered with available methods?*

• Feasibility
  *Can the question be answered with available time, resources, and expertise?*
Types of Research Questions

- What is the relative importance of less pain versus heart-attack risk?
- What is the money-equivalent value (WTP) of an effective treatment for treatment-resistant depression?
- How do preferences vary between patients at earlier and later stages of MS progression?
- How do patient-weighted EQ-5D scores differ from conventional scores?
- What is the possible uptake of a new weight-loss device?
- How adherent are patients likely to be with a new injection technology?
What Evidence is Needed?

• Evidence should answer the research question of interest

• Be as specific as possible
  – May require formally posing a hypothesis or hypotheses
  – Describe preference information needed to evaluate and test hypotheses
<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Implied evidence need</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducing severity of dyspnea from moderate to severe is more important than a 3% chance of pneumonia.</td>
<td>• Importance of symptom improvement relative to adverse-event risk</td>
</tr>
<tr>
<td>More than 75% of patients would accept a 5% increase in the chance of bleeding to improve physical functioning.</td>
<td>• Some measure of dispersion for importance of symptom improvement relative to adverse-event risk</td>
</tr>
<tr>
<td>• Individual or group-specific relative importance of symptom improvement relative to adverse-event risk</td>
<td></td>
</tr>
<tr>
<td>Improvements in mental functioning are more important than improvements in physical functioning.</td>
<td>• Importance ranking of treatment benefits</td>
</tr>
</tbody>
</table>
ISPOR Checklist for Stated-Preference Applications in Medicine

1. Research question
2. Attributes and levels
3. Construction of tasks
4. Preference elicitation
5. Instrument design
6. Experimental design
7. Data collection
8. Statistical analyses
9. Results and conclusions
10. Study presentation
Attribute-Selection Guidelines

• Attributes are generic features (vehicle type, color)
• Levels are variations within each feature (car/bus, green/red)

• Treatment attributes:
  – Clinically relevant and salient to respondents
  – Must be outside respondents’ control (physician choice)
  – Must vary independently (pain + ADL?)

• Levels:
  – Ranges wide enough to encourage tradeoffs
  – Include values observed or expected in clinical evidence
Identifying and choosing attributes

Top-down approach: trial data

Bottom-up approach: importance to patients

- Trial end points and AEs
- PRO items
- Literature reviews
- Social media
- Qualitative interviews
- Focus groups
- Expert opinion
- Patient advocacy groups
Choosing Levels

• Appropriate range of levels
  – Relevant clinical range
  – Range over which subjects are willing to accept tradeoffs

• Appropriate number of levels
  – Usually 2-5
  – Numeric
    • linear: 2 levels
    • quadratic: 3 levels
  – Categorical: relevant number of categories
### Example Attribute Table: Weight-Loss Devices

<table>
<thead>
<tr>
<th>Attribute Label</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average amount of weight loss (in lbs, based on reported weight)</td>
<td>- 5%</td>
</tr>
<tr>
<td></td>
<td>- 10%</td>
</tr>
<tr>
<td></td>
<td>- 20%</td>
</tr>
<tr>
<td></td>
<td>- 30%</td>
</tr>
<tr>
<td>On average, how long the weight loss lasts</td>
<td>- 6 months</td>
</tr>
<tr>
<td></td>
<td>- 1 year</td>
</tr>
<tr>
<td></td>
<td>- 5 years</td>
</tr>
<tr>
<td>Type of operation</td>
<td>- Endoscopic surgery</td>
</tr>
<tr>
<td></td>
<td>- Laparoscopic surgery</td>
</tr>
<tr>
<td></td>
<td>- Open Surgery</td>
</tr>
<tr>
<td>On average, how long side effects last that limit daily activities</td>
<td>- None</td>
</tr>
<tr>
<td>several times a month.</td>
<td>- 1 month</td>
</tr>
<tr>
<td></td>
<td>- 1 year</td>
</tr>
<tr>
<td></td>
<td>- 5 years</td>
</tr>
<tr>
<td>Chance of side effects requiring hospitalization</td>
<td>- None</td>
</tr>
<tr>
<td></td>
<td>- 5% chance of going to hospital with no surgery</td>
</tr>
<tr>
<td></td>
<td>- 20% chance of going to hospital with no surgery</td>
</tr>
<tr>
<td></td>
<td>- 5% chance of going to hospital for surgery</td>
</tr>
<tr>
<td>Chance of dying from getting weight-loss device</td>
<td>- 0%</td>
</tr>
<tr>
<td></td>
<td>- 1% (10 out of 1000)</td>
</tr>
<tr>
<td></td>
<td>- 3% (20 out of 1000)</td>
</tr>
<tr>
<td></td>
<td>- 5% (50 out of 1000) or 8% (80 out of 1000)</td>
</tr>
<tr>
<td></td>
<td>- 10% (100 out of 1000) or 15% (150 out of 1000)</td>
</tr>
</tbody>
</table>

1. Research question

2. Attributes and levels

3. Construction of tasks

4. Preference elicitation

5. Instrument design

6. Experimental design

7. Data collection

8. Statistical analyses

9. Results and conclusions

10. Study presentation
Construction of choice questions

• How many alternatives?

• Include opt-out or status-quo alternative?

• Unlabeled or labeled alternatives?

• How much information in the attribute labels?

• Decision frame: What motivates “Which alternative would you choose?”
How many questions?

• Need to consider complexity
• Number of attributes
  – Depends on attribute complexity
  – Rule of thumb: no more than 8
• Number of alternatives
  – Most studies use 2-3
• Number of choice questions per respondent
  – 5 to 10 common
Survey Instrument Components

- Explanation of study / Consent
- Attribute descriptions
- Risk tutorial
- Practice choice questions
- Comprehension tests
- Choice questions
- Background questions
Describe the Attributes

- Preferences of “well-informed” respondents
- Sufficient information for respondents to evaluate trade-off questions accurately—including something does not ensure respondents will understand or consider it
- Clinically accurate, but written at low reading level
- Cognitive overload—more information is not necessarily better
Ability to do daily activities

Health problems can make people feel tired or out of breath as they do their usual daily activities. Later in the survey, we will use the pictures below to describe three levels of severity of health problems. Please look at the table below to learn how such problems affect people’s ability to do daily activities.

Later in this survey we will use these descriptions to help you think about treatments that could help improve people’s ability to do daily activities.

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climbing stairs:</td>
<td>Not able to climb stairs</td>
<td>Must stop to rest when climbing stairs</td>
<td>No need to stop, but must climb slowly</td>
</tr>
<tr>
<td>Housework:</td>
<td>Cannot do housework</td>
<td>Must stop to rest when doing light housework</td>
<td>No need to stop when doing light housework</td>
</tr>
<tr>
<td>Resting:</td>
<td>Feel tired and out of breath when resting</td>
<td>Feel comfortable when resting</td>
<td>Feel comfortable when resting</td>
</tr>
</tbody>
</table>
Check for Comprehension

Use follow-up questions that require respondents to re-read the definition if necessary

**PRO type:**

Based on the descriptions above, which level best describes your ability to do daily activities today?

- [ ] Able to climb stairs without stopping, but slowly
- [ ] Able to climb stairs, but would have to stop to rest
- [ ] Cannot climb stairs

**Quiz type:**

Based on the descriptions above, how easily could people climb stairs if their limitations were level (2)?

- [ ] Able to climb stairs without stopping, but slowly
- [ ] Able to climb stairs, but would have to stop to rest
- [ ] Cannot climb stairs
Barriers to Valid Preference Data

• Risk and numeracy

• Preparation for choice questions
  – Decision frame
  – Practice with choice task

• Dealing with hypothetical bias

• Use of pretest interviews
Risk Attributes

• Uncertainty is a fact of life in health care
• Most people are averse to risk
• Risk attributes significantly increase choice difficulty
• Challenge of low numeracy levels in patient populations
• Need for simple risk tutorial
Example Risk Tutorial

Doctors do not know who will have a serious problem. However, based on their experience with large numbers of patients, doctors know how many people have had a serious problem after receiving a device. That information can help you think about your own chance of having a serious problem.

We will use some pictures to help you think about the chance of having a serious problem. The box below has 100 figures. Each figure represents a person who has received a device.

The blue figures show the number of people out of 100 who had a serious problem because of the device. The gray figures show the number of people out of 100 who did not have a serious problem.

In this example:
4 figures are blue. That means 4 people out of 100 (4%) who received a device had a serious problem.
96 of the figures are gray. That means 96 people out of 100 (96%) who received a device did not have a serious problem.
Check for Comprehension

Use follow-up questions to test understanding of risk graphics

In this example, how many people will get the severe side effect with **Device A**?

- 5 out of 100 (5%)
- 8 out of 100 (8%)
- 95 out of 100 (95%)
- 92 out of 100 (92%)

[If not 8 out of 100 (8%)]

The correct answer is 8 out of 100. Remember that the gray figures indicate people who will **not** get a severe side effect. The purple figures indicate people who will **get** a severe side effect. There are eight purple figures in Treatment A, indicating 8 out of 100 people (8%) will get a severe side effect.
Suppose that you are not able to do most of your daily activities because of your health problems. Your doctor suggests you consider two devices (Device A and Device B). These devices have had different effects for the typical patient.

You also could decide not to have either device (No Device). If you choose No Device, you would not be able to do most of your daily activities, but you will have no risk of serious side effects or risk of death from getting a device.
### Example Practice Question

<table>
<thead>
<tr>
<th>Limitations in daily activities</th>
<th>No Device</th>
<th>Device A</th>
<th>Device B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stairs:</td>
<td>Cannot</td>
<td>Stop to rest</td>
<td>Climb slowly</td>
</tr>
<tr>
<td>Housework:</td>
<td>Cannot</td>
<td>Stop to rest</td>
<td>No stop to rest</td>
</tr>
<tr>
<td>Resting:</td>
<td>Uncomfortable</td>
<td>Comfortable</td>
<td>Comfortable</td>
</tr>
</tbody>
</table>

| How many people died from getting the device | None | 1 out of 100 (1%) | 3 out of 100 (3%) |

| Which option would you choose? | | | |
### Building up to full choice question

<table>
<thead>
<tr>
<th>Limitations in daily activities</th>
<th>No Device</th>
<th>Device A</th>
<th>Device B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stairs</td>
<td>Cannot</td>
<td>Climb slowly</td>
<td>Cannot</td>
</tr>
<tr>
<td>Housework</td>
<td>Cannot</td>
<td>No stop to rest</td>
<td>Cannot</td>
</tr>
<tr>
<td>Resting</td>
<td>Uncomfortable</td>
<td>Comfortable</td>
<td>Uncomfortable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospital stays in 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 stays</td>
</tr>
<tr>
<td>4 stays</td>
</tr>
<tr>
<td>2 stays</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How many people had a serious stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
</tr>
<tr>
<td>3 out of 100 (3%)</td>
</tr>
<tr>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How many people died from getting the device</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
</tr>
<tr>
<td>10 out of 100 (10%)</td>
</tr>
<tr>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Which option would you choose?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
To make our study a success, we need your help with a problem we have in studies like this. Because our participants do not actually have to live with the results of the treatment they select, they often do not think carefully about what they would do if they really had to choose.

If you do not pay attention to the information shown in each question as you would in real life, we will not get a true measure of how important various treatment benefits and risks actually are to people like you.

We need your thoughtful answers to help us understand how you feel about possible medical devices.
Face-to-Face Pretest Interviews

• 10-12 respondents similar to target population
• “Think-aloud” protocol
• Probes
  • “Why did you choose A?”
  • “How could we make that clearer?”
  • “What was your reaction to ...?”
  • “If [level x] in alternative A were [twice/half] as large, would you switch from B to A?”
• Refine as necessary (not collecting data, but refining the instrument)
1. Research question

2. Attributes and levels

3. Construction of tasks

4. Preference elicitation

5. Instrument design

6. Experimental design

7. Data collection

8. Statistical analyses

9. Results and conclusions

10. Study presentation
# Experimental Design 101: Soup Recipes

## Attributes

<table>
<thead>
<tr>
<th>Levels</th>
<th>MEAT</th>
<th>NOODLES</th>
<th>VEGETABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Beef</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

## Soups

<table>
<thead>
<tr>
<th>Soups</th>
<th>MEAT</th>
<th>NOODLES</th>
<th>VEGETABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Chicken</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Chicken</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Chicken</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Chicken</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Beef</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Beef</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Beef</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>Beef</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
Most DCEs too large to use full factorials

FOUR ATTRIBUTES:

A: 4 levels
B: 3 levels
C: 5 levels
D: 2 levels

FULL FACTORIAL

= 4 \times 3 \times 5 \times 2 = 120 \text{ profiles}

7,140

2-alternative questions
Partial-Factorial Design Software

- SAS experimental-design macros
- Sawtooth Software
- Ngene
Conflicting Objectives

• **Reliable subject response:** requires simple designs

• **Granular data:** requires high-resolution designs

• **A lot about a little versus a little about a lot**
Practical Design Considerations

• Examine design for implausible combinations

• Examine combinations for dominated alternatives (perhaps used as a test)

• Block total design into individual versions of 5-10 questions

• Check versions for approximate level balance
1. Research question
2. Attributes and levels
3. Construction of tasks
4. Preference elicitation
5. Experimental design
6. Instrument design
7. Data collection
8. Statistical analyses
9. Results and conclusions
10. Study presentation

Topics for Webinar III
Reed Johnson  
reed.johnson@duke.edu  
919 668 1075

Shelby Reed  
shelby.reed@duke.edu  
919 668 8991

Juan Marcos Gonzalez  
jm.gonzalez@duke.edu  
919 668 5157
Join us for the final session

- April 19 - Session 3: Example applications and lessons learned—analysis and reporting

- Recordings will be available on http://mdic.org/mdicx