Data Collection and Analysis in Evaluation

February 2019

Alberta Health Services

Table of contents

Project team

Project sponsor
Neha Batra-Garga, Manager
Knowledge Exchange, Provincial Addiction and Mental Health

Prepared by
Michelle Chan, Senior Research Officer
Knowledge Exchange, Provincial Addiction and Mental Health

Contact

For more information, contact:
Knowledge Exchange, Provincial Addiction and Mental Health
Alberta Health Services
amh.knowledgeexchange@ahs.ca

For citation purposes, use the following format:
Alberta Health Services. (2019). *Data Collection and Analysis in Evaluation*. Edmonton, AB: Author.

Data collection

Data collection is the systematic process of gathering information from a variety of sources to answer your evaluation questions. This may include sources like surveys, interviews, and focus groups.

Sampling

The goal of sampling is to obtain participants that are representative of the study population. Since it is often too time-consuming, costly, and impractical to survey everyone, you can use information from a subset of the population to draw conclusions about the larger population (Ponto, 2015). Population lists may come from service registration lists, telephone number listings, business registries, or random digit dialing. You will need to be aware of privacy laws when using personal information for your research or evaluation.

Common types of sampling:

- **Systematic:** Choose every *n*th (for example, 10th) name from a list.
- **Simple random:** Everyone has an equal probability of being chosen.
- **Purposive:** Participants are pre-selected (such as key informants or hard-to-find populations).
- **Snowball:** Participants inform the researcher of others to contact.
- **Convenience/haphazard:** Anyone who is eligible can participate.

Complex sampling:

- **Stratified:** Participants are chosen based on pre-determined groupings (such as gender or age groups). This is sometimes necessary to ensure your sample is representative of the study population.
- **Quota:** Specific targets for the number of participants that you want to sample are set (e.g., 100 males and 100 females). Sampling is done after you've reached each target.
- **Two-stage cluster:** A sample of clusters (for example, colleges) is identified, then a sample within each of those clusters (for example, students) is chosen (Ray, 2011).

Sometimes sampling techniques are combined. For example, systematic sampling may begin with a simple random number or a stratified sample.

In general, survey sampling involves a degree of error (e.g., ± 5%). The degree of error is the amount of error that you can tolerate. To calculate the sample size you need, you will need to knowing the size of the population and the acceptable degree of error (Ray, 2011).* It is sometimes necessary to know the size of the subgroups you want to analyze (such as gender

or age) in order to determine the number you need to survey within each subgroup. There are online resources to assist with calculating sample sizes, such as this Sample Size Calculator. You also need to consider your expected response rate. For example, if you expect 50% of those you contact to agree to participate, your total sample will need to be twice the number needed to attain the accepted degree of error.

Recruiting

Recruiting participants is a crucial step in data collection. If too few people agree to participate, you may introduce *non-response bias.* This means there may be something unique about those who did not respond or the answers they would have provided. As a result, you may not be able to generalize your findings to the larger population.

Strategies to help recruit participants:

- Keep the survey as short as possible. An ideal survey takes less than 10 minutes to complete. Tell participants that the survey is short, as this may increase response rates (Public Works and Government Services Canada, 2007).
- Limit telephone surveys to less than 20 minutes. In-person interviews and focus groups can be longer (an hour or more) (Newman, 2010).
- If appropriate, share the project on the Internet, in newspapers, in newsletters, or other forms of media.
- If possible, include incentives to reimburse participants for their time or travel expenses. This may be a payment, gift card, or a chance to win a prize (French, 2012). Generally, these incentives do not need to be large to be effective. Also, if they are more than what would be considered “reasonable reimbursement,” they may be considered coercive.
- Send up to two reminders after participants receive a mail or online survey.

Note that some agencies or government ministries have a communications department that you must consult before recruiting participants. There may also be policies regarding acceptable and appropriate incentives.

Designing data collection instruments

Principles of good question writing

Below are some key principles to help ensure that your questions will be understood and answered as intended.

Write from the perspective of the participant.
Ensure that the readability and grade level of the questions is appropriate.

- Microsoft Word has a built-in function that will give you the Flesch-Kincaid Grade Level for any sample of writing.

- A grade seven to grade eight reading level is recommended for adults (Baxter & Babbie, 2003).
- You will want to adjust the grade level if your participants may have lower literacy or difficulties with comprehension. For example, children or those whose first language is not English.

Be clear and concise.

- The clearer you make the question, the more reliable the data will be. For example, ask "What is your annual income before taxes?" instead of "What is your income?"
- Avoid abbreviations, slang, and jargon.
- Avoid words that could be open to interpretation. For example, instead of "regularly" use "more than more a week"
- Avoid double-barreled questions. For example, "Do you own a house and vehicle?" Separate this into "Do you own a house?" and "Do you own a vehicle?"

Avoid asking about future intentions.

- In general, people are poor predictors of their behaviour (Ajzen, 2011).

Use neutral language.

- Emotional or leading language can produce biased responses. For example, the question "Do you think hitting children is wrong?" tends to provoke the answer "yes" because the phrasing suggests that hitting children is wrong (Ruane, 2016).

Avoid overlapping or unbalanced response options.

- Overlapping options: 1-5, 5-10, 10-15
- Unbalanced response options: Excellent, Good, Poor. This scale has an uneven number of positive and negative response options.

Avoid asking unnecessary questions.

- There are many questions that will give you information that would be *nice* to know. You should only collect data that you *need* to know, and is linked to one of your evaluation questions.

Using existing instruments

You may find an existing data collection instrument that addresses your evaluation question. Before using it, you should ensure that it is:

- Valid: Measures what it intends to measure
- Reliable: Produces consistent results

If it meets your requirements, you will need to contact the author for permission to use the instrument.

Survey design

Surveys are often used to collect information because:

- They are relatively cost-effective and practical for a large sample.
- You can use well-defined structured questions (such as multiple choice).

Surveys typically ask questions related to knowledge, beliefs, attitudes, behaviours, and characteristics (Baxter & Babbie, 2003).

Things to consider when designing a survey (French, 2012; Newman, 2010):

- Keep your survey as short as possible. An ideal survey takes less than 10 minutes to complete.
- Provide an introductory statement explaining the purpose of the survey and any instructions.
- Begin with questions that are easy to answer, non-threatening, and non-sensitive.
- Group questions on the same topic together and provide an introductory statement for new topics.
- Pilot your survey before using it with an entire sample (see below).

Writing survey questions

Survey questions typically fall within one of four categories:

1. **Ordinal questions** have responses that are fixed on a continuum or scale and the order of responses is meaningful. For example, they may measure levels of satisfaction, agreement, knowledge, awareness, or frequency. The order of responses is *relative*, not exact. For example, you can't quantify the exact difference between "very satisfied" and "satisfied". Options for ordinal responses include:
    - Rating questions (such as a Likert scale)
    - Ranking questions (the responses are numbered in order of importance)

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

Very Unsatisfied	1	2	3	4	5	Very satisfied
	○	○	○	○	○	

2. **Categorical (nominal) questions** have multiple response categories, but there is no natural order or ranking of the categories. For example, gender, blood type, types of books, yes/no.

3. **Interval questions** are similar to ordinal questions but there is a meaningful mathematical difference between responses. For example, age, IQ, or temperature. The difference between 20 and 16 degrees is the same magnitude as the difference between 28 and 24 degrees.

4. **Open-ended questions** allow for free text that is not structured. For example, narrative responses (“What is the solution to global warming?”) or if there is a wide range of potential responses (“What is your favourite food?") (Couper et al., 2011).

Tips for writing survey questions:

- Using scales allows for standard responses across surveys and makes for faster and easier analysis for large samples.
- Be sure that the response options do not overlap with one another. For example, “grocery store” and “supermarket” (French, 2012).
- Make it clear on the survey when participants can choose multiple responses to a question.

Interview design

Interviews are often used to collect information when:

- There is a smaller sample size (for example, key informants).
- In-depth information is required.

Interviews can be:

- **Structured:** There is a fixed set of questions that are asked in a specific order.
- **Semi-structured:** There are fixed questions, but new questions can be introduced if necessary.
- **Unstructured:** There may be a goal for the interview, but there is not a formal interview guide.

It is common for interview questions to be open-ended. Interviewers may also use probes to gather additional information, or clarify a question if the participant does not understand it.

Interviewers should remain neutral during the interview to avoid introducing bias. Interviewers may influence responses if they have expectations about the answers or if they probe inadequately or inappropriately. Interviewers may also influence the way participants respond by their tone, attitude, or reactions to answers (Newman, 2010).

Focus group design

Focus groups are often used to collect information when:

- There is a smaller sample size (this small sample may be drawn from the larger sample).
- Discussion or brainstorming is required.

Focus groups are similar to interviews in that they often include:

- Semi-structured guides
- Open-ended questions
- Probing questions

As with interviews, it is important for the focus group moderator to avoid introducing bias. It is important to stress that there are no incorrect answers and to ensure that all participants have an opportunity to talk.

Pilot testing

You will likely want to have your new surveys, interview guides, and focus group guides reviewed and tested before using them with the study sample. This can help ensure questions are clear to both the interviewers and respondents. The test can be used to identify errors in wording and skip patterns (questions associated with conditional responses in a survey), and can indicate how long the surveys or interviews will take (French, 2012).

Things to consider for pilot testing:

- The pilot participants should have similar characteristics to the larger sample.
- Ask the pilot participants questions after completing the survey. For example: Was there was anything they did not understand? Did they think some issues were overlooked? Did they think it was too long?
- Ask the interviewers about any potential issues.
- Pilot data can be combined with the remaining survey data if no significant changes are made after the pilot.

Ethical considerations in data collection

Participants are often required to provide voluntary, informed consent prior to participating in any research or evaluation project. In some situations, informed consent is not necessary. For example, an evaluation that uses observation methods in a natural setting. In these cases, a sign that alerts individuals that a study is being conducted may be sufficient. Even then, it is important to provide individuals with information about the study (evaluation purpose, methods, how to opt-out, contact information) in case they would like to know more about it.

Typical statements within a consent form includes:

- Purpose, process, and duration of the project
- Any potential risks to the participant
- Guarantee of anonymity and confidentiality
- Researcher's information and contact information
- Indication that their participation is voluntary and they can withdraw at any time
- Information about any benefits and compensation
- Information about how to access results (Newman, 2010)

Data should be collected in a way that minimizes the burden placed on participants. This can be done by pilot testing the data collection instrument and ensuring that it is not too time-consuming, difficult, or emotionally stressful.

The participants' information should be kept confidential both during and after data collection. Collected data should be kept in a secure place (such as inside a secure building, using password protection and/or encryption on electronic files or memory devices).

Data analysis

Basic quantitative analysis techniques

Quantitative data analysis uses statistics to provide quantifiable, objective, and easy to interpret results from numeric data.

Descriptive statistics can be used to describe or summarize your data:

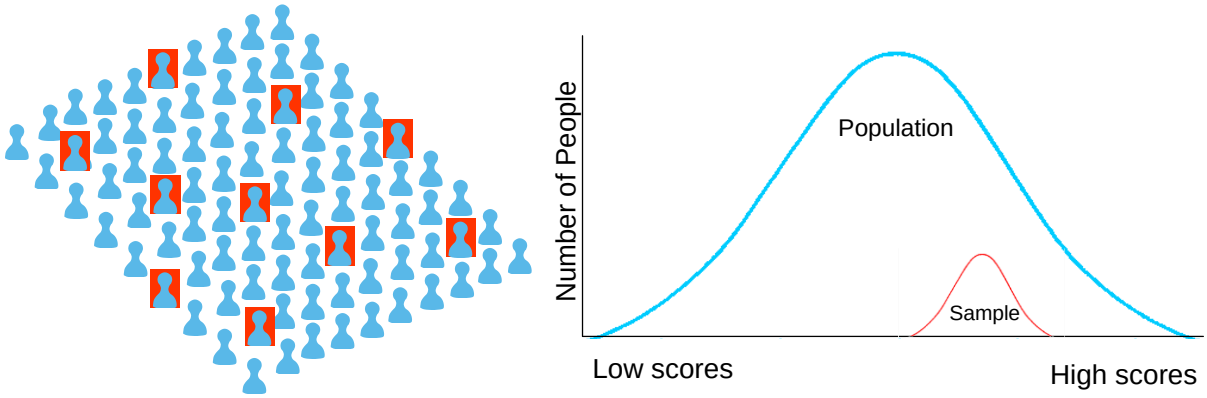
- N (number of participants)
- Mean (average)
- Median (middle value)
- Mode (value that occurs the most often)
- Standard deviation (measure of variability)
- Minimum and maximum
- Crosstab (how many of a certain category were answered in a certain way)
- Frequencies (number of times you got a certain response)

Inferential statistics can be used to understand the relationships between variables and to make generalizations about the larger population based on your sample. Examples include:

- Correlation (how much two variables are related)
- Regression (how much one variable changes with any given change of another variable)
- *t*-test (compares the means of two groups)
- Analysis of variance (compares the means of more than two groups)

Tests for significance

- A null hypothesis is a statement that proposes there is no statistical significance between two variables.
- An alternative hypothesis is one that states that there *is* a statistically significant relationship between two variables.
- A result is called **statistically significant** if it is unlikely to have occurred by chance.
- Your data contains is a sample (a small group) who are meant to represent everyone in a population (a larger group). Since you randomly select people, their scores will never be *exactly* what the whole population's score would be. Significance testing gives you an idea of the probability of your result, other than by chance alone.



How probable does a result have to be to be considered significant?

- A cut-off point is used to determine whether a result is statistically significant.
- In general, a *p*-value of 0.05 or lower is considered statistically significant.
    - $p < .05$; less than 5% probability of obtaining your result without a true difference
    - $p < .01$; less than 1% probability of obtaining your result without a true difference
- A result with $p > .05$ (or your determined cut-off point) is not statistically significant, even if the *p*-value is .051 (some people might call this "marginally significant" or say there was a trend in a certain direction).

Cautions:

- The term *significant* does not imply *important* or *meaningful*, as it does in everyday speech. In other words, you can have a statistically significant difference between two groups, but the actual difference is not large enough to change your program because of it.
- Statistical significance does not mean that your theory is true. Correlation does not equal causation!
- The *p*-value is not an indicator of the strength of the relationship between two variables.

Basic qualitative analysis techniques

Qualitative data analysis involves recognizing patterns in your data (text, images, audio, or video) and turning patterns into meaningful ideas. It involves developing schemes to organize and manage your data through:

- Coding or categorizing information within each case (for example, each interview transcript).
- Identifying themes across all cases.

Coding qualitative data

Coding is often the first step in qualitative data analysis. It is the process of organizing and sorting your raw data (e.g., interview transcript), and converting it into usable data through the identification of themes and patterns. It is also often used with other forms of data analysis, such as thematic analysis.

Coding is typically a two-step process:

1. **Developing codes/categories**
    - Codes can be generated from your data collection tool and/or data.
    - Generating codes from your data will typically involve reading through your data and making notes on the data you are analyzing.
    - Consider both recurring patterns and deviant cases.
    - If more than one person is involved in your evaluation, it is useful to develop codes independently and then compare and combine them.
    - Develop a coding sheet or form (a guide to remind you about what information to look for).

2. **Coding data**
    - This is a more systematic review of your data.
    - It involves going through your data and filling out a coding sheet for each case (for example, each interview).
    - Coding can be an iterative process. As you are coding your data, you may find that you need to go back and update or revise your coding sheet.
    - It is often helpful to identify quotes or examples to illustrate the idea the code is meant to capture.
    - You can use your codes to generate basic descriptive quantitative data (such as counting the frequency of specific codes).

Thematic analysis

Thematic analysis identifies broad themes or more complex patterns in your data. It involves:

- Recognizing relationships and links within your data or between your codes/clusters of codes.
- Identifying processes and outcomes.
- Determining substantive significance (how important, meaningful, or potentially useful are your findings?).
- Interpreting the meaning.

This is done by looking *across cases*, whereas coding is typically done on a case-by-case basis.

Some tips for analyzing themes:

- Your coded data can help identify themes, but you may also have to revisit your data to fully develop your themes.
- Themes are identified from your data but may be influenced by your evaluation objectives.
- Similar to coding, it is helpful to identify quotes or examples that illustrate your themes.
- When identifying themes, you may want to use diagrams or matrices to help understand the links and relationships in your data. Some useful tools include concept maps, process maps, or taxonomies.

References

Azjen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health, 26*, 1113-1127.

Baxter, L. A. & Babbie, E. R. (2003). Survey research. *The basics of communication research.* Boston, MA: Cengage Learning.

Couper, M. P., Kennedy, C., Conrad, F. G., & Tourangeau, R. (2011). Designing input fields for non-narrative open-ended responses in web survey. *Journal of Official Statistics, 27*, 65-85.

French, J. (2012). Designing and using surveys as research and evaluation tools. *Journal of Medical Imaging and Radiation Sciences, 43,* 187-192.

Newman, W. L. (2010). *Social research methods: qualitative and quantitative approaches* (7th ed.). Boston, MA: Pearson.

Ponto, J. (2015). Understanding and evaluating survey research. *Journal of the Advanced Practitioner in Oncology, 6*, 168-171.

Public Works and Government Services Canada. (2007). Best practices in public opinion research: Improving respondent cooperation for telephone surveys. Ottawa, ON: Author.

Ray, W. J. (2011). Questionnaires, survey research, and sampling. *Methods toward a science of behavior and experience* (10th ed.). Boston, MA: Cengage Learning.

Ruane, J. M. (2016). The questionnaire: Would you mind taking the time to answer a few questions? *Introducing social research methods: Essentials for getting the edge.* Hoboken, NJ: Wiley-Blackwell.