

# Types of Data Used for Impact Evaluation

OPRE Report # 2021-117



Mark Courtney

University of Chicago School of Social Service Administration

# Agenda

- Part 1: Introduction to primary and secondary data sources for use in evaluating interventions
  - Teams consider options for using these data sources for their evaluations
- Part 2: Relative strengths and limitations of the major data sources used in evaluating interventions
  - Teams reconsider their options in light of characteristics of potentially available data
  - Q&A based on team priorities

# Learning Objectives

- Participants will be able to describe the major data sources used in program evaluations.
- Participants will be able to describe the relative strengths and limitations of the major data sources for program evaluation purposes.
- Participants will be able to apply their knowledge of data sources used in evaluation to assess the relative merits of data sources for an evaluation they are involved in planning.

# Data Collection Should Be Driven by Your Theory of Change. For Example...

- What information will you need to assess whether your intervention is serving the target population?
- What information will you need to assess whether the intervention had an impact on the kinds of help the target population received?
- What information will you need to know whether outcomes differed between your intervention and control or comparison groups?

*Keep your theory of change in mind throughout design of data collection!*

# Two Types of Evaluation Data: Primary and Secondary

*Primary data* are collected directly from a data source (e.g., an individual involved in an impact evaluation) for the purposes of an evaluation without relying on already available data sources.

*Secondary data* have been collected for a purpose other than the evaluation in question but are available for use in an evaluation.

# Secondary Data Useful for Capturing Outcomes of an Intervention Include Administrative Data

- Administrative data for measuring outcomes experienced by individuals participating in an evaluation include the following:
  - **From child welfare agencies:** demographic information; child maltreatment and foster care histories; caseworker assessments of parent and child functioning; foster care placement mobility; permanency outcomes; child maltreatment reports.
  - **From other governmental and nongovernmental sources:** information about services received from other public agencies (e.g., Medicaid; behavioral health systems; receipt of public aid (e.g., TANF, SNAP, and SSI); outcomes experienced by an evaluation population (e.g., juvenile justice or adult criminal justice system involvement; secondary and postsecondary educational outcomes (data from the public education systems and/or the National Student Clearinghouse); earnings based on data from state Unemployment Insurance systems.

# Secondary Data Sources Can Also Be Useful for Measuring the Context of Service Delivery in Impact Evaluation

- Administrative data and data from other publicly available sources can help measure the *context of service delivery*.
  - Service context includes anything that could influence the impact of an intervention on the target population (e.g., costs of basic necessities including housing; labor market; availability of services that could substitute for or affect the impact of an intervention being evaluated).
  - Sources of this information include the Comprehensive Child Welfare Information System (CCWIS), US Census Bureau, and Bureau of Labor Statistics.
- This can be particularly important for evaluations that use quasi-experimental evaluation designs.

# What Kinds of Secondary Data Collection Could Be Useful for Your Evaluation?

- Take some time with your team to discuss these questions (and take notes!):
  - Could you use administrative data to measure the outcomes that are the focus of your evaluation?
  - Who would you need to get those data from, and what would be required for you to access those data?



# Surveys Are One Primary Data Source for Collecting Information as Part of an Impact Evaluation

- A survey collects information directly from individuals by asking them questions.
- Surveys are useful for collecting information individuals can reliably provide themselves.
- Surveys are not as effective at collecting information when social desirability of individuals' responses is a concern.
  - *Social desirability* is the tendency of some participants in a study to report an answer in a way they deem to be more socially acceptable than would be their “true” answer. They do this to project a favorable image of themselves.

# Surveys Can Be Carried Out Using Various Methods

- Examples of survey methods:
  - Mailed questionnaires
  - In-person interviews
  - Phone interviews
  - Web-based questionnaires
  - Text messaging

# Assessments Are Another Primary Data Source for Collecting Information as Part of an Impact Evaluation

- An assessment uses specialized methods (e.g., trained observation; psychological testing; medical testing) to gather information about individuals that they are unlikely to provide themselves.
  - Assessments often require special training of the assessor.
  - Some assessments can be included in surveys whereas others cannot.

# What Kinds of Primary Data Collection Could Be Useful for Your Evaluation?

- Take some time with your team to discuss these questions, based on your current understanding of the kinds of information you need for your evaluation (and take notes!):
  - Will you want to include surveys? Of whom? What survey methods make the most sense?
  - Would your evaluation benefit from collecting information through assessments?

# Issues to Consider When Weighing the Value of Primary and Secondary Data for Evaluation

- Validity of the data source
- Reliability of the data source
- Feasibility of data collection
- Cost of data collection

# Issues to Consider When Weighing the Value of Primary and Secondary Data for Evaluation: Validity of the Data

- **Validity:** How well the data measures what it is intended to measure and how well the data provide strong evidence for the conclusions made by an evaluation.
- Data sources vary in their validity:
  - Social desirability: Administrative data on parental involvement with child welfare services are generally more valid than data collected from surveys of parents or children.
  - Differences in attrition between groups: administrative data.

# Issues to Consider When Weighing the Value of Primary and Secondary Data for Evaluation: Reliability of the Data

- **Reliability:** The extent to which a measurement, when repeatedly applied to a given situation consistently, produces the same results if the situation does not change between the uses of the measure.
- Reliability can refer to stability of the measurement over time or to the consistency of the measurement from place to place.
  - Reliability over time may suffer if respondents are annoyed by a measure and find ways to shorten the time they take to respond to the measure (e.g., over time they come to always answer “no” to questions about mental health symptoms).
  - Reliability of measurement between sites where measures are collected may suffer if the data collection practices vary between sites, such as when casework practice around data entry varies between counties.

# Issues to Consider When Weighing the Value of Primary and Secondary Data for Evaluation: Feasibility

- **Feasibility:** How easy the data are to collect
- Aspects of feasibility of data collection include the following:
  - Ease of engagement of evaluation participants
  - Burden of data collection on service providers and others affected by evaluation activities
  - Ease of access to administrative data
  - Costs of data collection



# For What Purposes are Primary And Secondary Data Sources Most Valid and Reliable?

- In comparison with administrative data, well-designed surveys can provide relatively detailed measures of a wide range of outcomes of interest to evaluators, particularly aspects well-being that are poorly captured by administrative data.
  - Prioritize using existing survey measures with well-established validity and reliability.
- Administrative data are most useful when capturing outcomes that informants are likely to have difficulty recalling accurately and outcomes where social desirability bias is a concern.
  - Missing data on outcomes because evaluation subjects move outside of the area covered by the administrative data can also limit the data's reliability and validity.

# What Affects the Feasibility of Using Survey Data for Evaluation?

- Feasibility of **using survey data for impact evaluation** involves complex trade-offs between the various methods of survey administration and the reliability and validity of the data collected.
  - **Ease of access to subjects** affects which types of surveys are feasible and the likely level of participant attrition during an evaluation.
    - Phone surveys require participants to have phones and reliable phone numbers.
    - Web-based surveys face similar access issues for evaluation participants.
    - Failure to assess evaluation participants' ability and inclination to participate in less engaging modes of survey administration (i.e., paper-and-pencil questionnaires, phone surveys, and web surveys) is likely to result in poor baseline response rates.
    - Attrition tends to be lower when using in-person engagement.

# What Affects the Feasibility of Using Survey Data for Evaluation? Cont'd.

- Reliable data on outcomes often requires **careful attention to the limits of some methods of survey data collection.**
  - Participants with limited reading ability may not be able to answer questions that require them to read questions and response options. This can rule out all modes of data collection except in-person or phone interviews.
  - Social desirability bias is likely to be greater when asking sensitive questions, particularly using in-person interview methods. Audio-CASI is preferable when outcomes of interest, like criminal justice system involvement or behavioral health problems, are of key interest.

# What Affects the Feasibility of Using Administrative Data for Evaluation?

- Feasibility of using administrative data for impact evaluation is primarily influenced by the availability and quality of data and ability of key evaluation partners to carry out tasks they will be responsible for.
  - The characteristics of local child welfare administrative data and the data management capacity of the child welfare agency influences the feasibility of using administrative data for evaluation purposes.
  - Relationships between the child welfare agency and other public institutions that have potential outcome data and the capacity of agencies to link data across data systems also influence the feasibility of using administrative data for evaluation purposes.

# What Are the Costs of Various Data Sources?

- **Survey data collection costs** vary widely depending on the characteristics of the evaluation population, the method of survey administration, and the costs of any assessments involved.
  - **Population characteristics:** Surveys of populations that are relatively easy to locate are less expensive on a per person basis than surveys of harder-to-locate populations. Evaluation populations that require extra effort to collect data from (e.g., those with disabilities, limited reading ability, or whose primary language is other than English) also increase the costs of survey data collection.
  - **Survey Method:** Mail-in surveys, phone surveys, and web- and text-based surveys are less expensive than in-person interviews. Adding CAPI and Audio-CASI increase the costs of data collection, but they can increase the reliability of data and reduce the costs of data management before analysis.
  - **Cost of Assessments:** Assessments can increase the costs of a survey marginally or quite substantially.

# What Are the Costs of Various Data Sources?

- **Administrative data costs** generally consist of the time devoted by public agency staff and evaluators to negotiating data-sharing agreements, generating data for the evaluation, and familiarizing evaluators with the strengths and limitations of the data.
  - Those costs can vary widely between evaluations depending on the experience of the collaborating agencies in using administrative data and linking data across systems (if called for) and the familiarity of evaluators with the administrative data involved.
  - The costs per study participant to obtain the measure of an outcome is much lower on average than would be the case for survey data on similar outcomes.

# Regroup and Q&A

- Discuss your evaluation plans with your team in light of the issues we just explored.
- 10 minutes in your group to identify two questions for Mark.
- Your questions can generally concern the uses of survey and administrative data for evaluation purposes, or be more specific to your project.
- Post your questions using the chat function ASAP.



This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Urban Institute et al. (2021). Slide Deck Session 12: Types of Data Used for Impact Evaluation - Child Welfare Evidence-Building Academy. OPRE Report 2021-117, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.