A GUIDE TO ACCESSING AND USING WORKFORCE DATA

Prepared by:
Correne Saunders
Valerie Benson

BOLD
THINKERS
DRIVING
REAL-WORLD
IMPACT

OVERVIEW

Data are a powerful tool for agencies and programs in their work to improve the lives of youth. Data can help document the need for a program or service, communicate how a population is faring, or assess a program's effect on the outcomes of those they serve.

This guide provides an overview of existing administrative data that can be used to measure youth workforce outcomes, including employment and educational outcomes. It also provides practical suggestions for accessing and using such data.

The guide begins with a discussion of the different types of data and how it is collected and then offers guidance for acquiring and understanding data from administrative sources. It then describes examples of administrative data sources common to measuring workforce outcomes, including state WIOA program data, state Unemployment Insurance (UI) earnings data, and National Student Clearinghouse college data.

HOW TO USE THIS GUIDE

We offer some helpful tips below on how to use the guide.

Click to navigate to different topics covered in the guide

Links to additional information are provided in these text boxes.

We denote the "pros" or advantages using this symbol.

Limitations or "cons" are denoted using this symbol.

Helpful tips are provided in these text boxes.

TYPES OF DATA COLLECTION

Program evaluation requires data to measure outcomes, and there are two basic ways to collect it.

PRIMARY DATA COLLECTION

The first is to collect data specifically tailored for the evaluation, usually by asking the participant to fill out forms or answer a survey at various points during the intervention. This is referred to as **primary data** collection.

Primary data can be very expensive and arduous to collect. The benefit is that the information is tailored for measuring the outcomes of a particular study and is usually easier to work with in terms of data quality. There is a large body of research and guidance for survey methodology that is beyond the scope of this guidebook.

OVERVIEW

HOW TO USE THIS GUIDE

TYPES OF DATA

GENERAL GUIDANCE
- Acquiring the Data
- Understanding the Data

EXAMPLES OF WORKFORCE PROGRAM DATA
- Program Data
- Earnings Data
- Education Data

SECONDARY DATA COLLECTION

The second kind of data collection is called **secondary data** collection. This refers to using databases that have already been created usually for purposes unrelated to the study itself, including census data, national survey data, and also what is referred to as "program" or "administrative" data.

Administrative data refers to data that caseworkers or administrators collect in the process of determining eligibility or tracking progress and compliance in a program like cash assistance, unemployment insurance benefits, or medical assistance.

Secondary data can sometimes be less expensive to acquire, but there can be immovable barriers to gaining access. It can also be difficult to retro-fit the information for a study that the data were not intended to address.

For instance, you may not be able to measure the primary outcomes of interest in the study, at least not in the way that you might have hoped, using the existing data.

GUIDANCE FOR ACQUIRING AND UNDERSTANDING ADMINISTRATIVE DATA

GUIDANCE FOR USING ADMINISTRATIVE DATA

Unlike the wide body of research that is available on survey methodology, very little is available in terms of guidance for using administrative data. In this section we provide guidance and things to consider when acquiring and using administrative data for your study, including:

ACQUIRING THE DATA

» Protection of Human Subjects
» Data Security
» Building Partnerships
» Timeline

UNDERSTANDING THE DATA

» Unit of Analysis
» Data Quality
» Data Documentation
» Context
» Archiving

ACQUIRING THE DATA:

PROTECTION OF HUMAN SUBJECTS

Program data are usually carefully guarded by the programs that own the data, with highly restricted access even if the data do not include direct PII (personally identifiable information), especially to non-governmental agencies or researchers. This is particularly true for vulnerable populations, including youth.

If you are interested in gaining access to administrative data for a particular program or in a particular jurisdiction, you should do some research on the applicable laws and regulations that govern data sharing.

You can find more information for understanding IRB requirements and whether your research requires review on the Department of Heath and Human Services website.

Research using administrative data may be considered human subjects research and therefore subject to review by the Internal Review Board of the state or grantee's home institution as well as any applicable federal and state laws and regulations.

ACQUIRING THE DATA:

DATA SECURITY

Data security is critical to protecting the confidentiality of the information acquired from administrative data and the privacy of individuals for whom you have data.

You should also build capacity for data security and familiarize yourself with national standards for data encryption, storage transfer, and destruction protocols. Be prepared to document your data security measures in a binding **Data Use Agreement (DUA)**.

See Harvard University's classification system for determining the sensitivity of data and corresponding data security requirements.

MIT provides examples of Nondisclosure agreements and DUAs.

For an example of an agreement with a state agency, see New York State Department of Labor's Data Sharing Agreement.

ACQUIRING THE DATA: BUILDING PARTNERSHIPS

Sharing administrative data comes with high stakes on the part of the agency that owns the data. It is important to build partnerships and be intentional to seek out ways your partnership can add value to the agency.

The better the partnership, the more likely you will receive valuable data documentation and programming support that will help you analyze and interpret the data.

Add value by providing rapid-cycle reporting based on the data you acquire in order to support their program management, or be available to provide brief analyses for particular management challenges or shifts.

ACQUIRING THE DATA: TIMELINE CONSIDERATIONS

The process for acquiring administrative data is lengthy, often requiring approvals from multiple parties. Start the process early in the design phase and prepare a contingency plan should data acquisition be delayed or if the data provided is not as expected and severely limits your analysis.

From the time of initial contact to obtaining the first data extract can range from 6 to 22 months.

Approval entities may not review data requests regularly. For example, IRBs may only meet quarterly.

Get a head start on the process by preparing a one-page summary of your project and include a description of how you will use the requested data, protocols in place to safeguard data, and how the protocols align with confidentiality requirements.

Data may available on a lagged basis for reporting purposes or to allow time for data entry. Once received, be sure to allow enough time for data cleaning and working with the data provider to understand and interpret the data.

UNDERSTANDING THE DATA: UNIT OF ANALYSIS

The data are sometimes collected at a caseload- or group-level instead of at the individual-level so it is important to know the structure of the data and how it may or may not correspond to the types of outcomes the study requires.

For example, group-level information might correspond to the costs and benefits of a particular program, or to the percent of a particular group or caseload that achieved a certain outcome.

Individual-level outcomes related to workforce initiatives include whether an individual found a job, achieved a diploma, or gained job skills.

When we rely on secondary (and particularly program) data, it may be necessary to confine the outcomes of interest to the data available.

Consider the content of program data very early on in the evaluation design process, and request a test file to help refine the data request and study design.

UNDERSTANDING THE DATA:

DATA QUALITY

Programs are typically not concerned with having perfect or clean data, especially if the data are not directly used to measure program performance or tied to administering monetary payments.

For example, it is not uncommon for administrative data to contain multiple records per individual, without a way to merge the information together, or for information to have typos or be outdated.

Think carefully about the nature of the information you are hoping to collect. Inquire about the quality of the measures early in the design process.

UNDERSTANDING THE DATA: DATA DOCUMENTATION

Documentation in the form of coding manuals or maps that can help you discern which variables to use and how, is worth pursuing but usually difficult to obtain.

Be aware of how current (or dated) the documentation is, and ask about "workarounds" that program administrators may use to accommodate changes in program rules over time that are not included in the original documentation.

Request a manual or materials used to train new program staff on data entry procedures. These materials better reflect current procedures for how data are collected.

Sample scenario: There is a new rule that requires participants to undergo drug screening, but nowhere in the current database to document compliance. A program administrator puts out guidance that caseworkers should use the "Birth Hospital" data field (which is no longer being used) to record compliance with a code of "Y" for compliance and "N" for non-compliance. This is called a workaround.

OVERVIEW

HOW TO USE THIS GUIDE

TYPES OF DATA

GENERAL GUIDANCE
- Acquiring the Data
- Understanding the Data

EXAMPLES OF WORKFORCE PROGRAM DATA
- Program Data
- Earnings Data
- Education Data

UNDERSTANDING THE DATA: CONTEXT

Give careful consideration when you are thinking about comparing or pooling data across sites or localities because context is important. There may be variables or measures that sound the same (i.e. “became employed”) but are understood differently in context (i.e., “became employed in a job that pays more than $10/hr” vs. “became employed at all”).

UNDERSTANDING THE DATA: ARCHIVING

Programs often have a schedule for archiving or expunging data over time, so it may be impossible or at least very difficult to access historical information.

COMMON SOURCES OF WORKFORCE ADMINISTRATIVE DATA

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

What is it?

The Department of Labor funds six core workforce programs, administered in states via local workforce development agencies. These programs are required to measure and report on their performance on individual and program outcomes, such as the number of individuals who completed credentialing programs and the number of individuals who obtained jobs after completing or exiting a workforce program.

The definition of specific performance measures for particular programs is currently being discussed and possibly modified, in order to meet requirements included in the Workforce Innovation and Opportunity Act (WIOA). The new measures will be released this summer and will likely require states to modify their own data collection systems in order to comply.

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

What can I measure using these data?

Based on the reporting requirements for workforce programs, the following may be measured using state administrative data for WIOA programs:

Service Receipt

- Engagement in career-related activities, such as completing a career exploration workshop, attending an occupational skills training class, participating in on-the-job training.
- Receipt of career guidance and counseling, supportive services, or referrals to jobs or additional training services.

Keep in mind, the availability of data to construct these measures varies by state.

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

What can I measure using these data?

Employment Outcomes:

- Employment outcomes, such as whether the youth entered employment and whether the youth retained that employment over consecutive quarters.
- 6-months' average earnings for youth retaining employment over consecutive quarters.

Educational Outcomes:

- Credential or degree attainment, such as high school diploma, GED, or certificate
- Literacy and numeracy gains for out-of-school youth who are basic skills deficient, such as whether functioning levels increased
- Enrollment in post-secondary education, advanced training, or occupational skills training program.

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

How can I gain access to these data?

Each state collects individual-level and program-level data for their workforce programs, and access to individual-level identifying information is restricted. Access to the data is controlled by states and local workforce agencies, and there are typically laws that govern the circumstances within which data may be shared. Contact the state workforce agency for these details.

Keep in mind that if data are available and the state is amenable to sharing the data for the purpose of a research study, a **Data Use Agreement** needs to be ratified by all involved parties.

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

What should I consider before making a data request?

Data are only available for individuals receiving services from local workforce agencies or participating in training funded by workforce programs. Data on longer-term outcomes related to college attainment or earning are not available.

There is a lag (usually six to nine months) between when wages are earned and when they are reported through different channels and then become available.

DOL has written a helpful document on the timing and availability of quarterly employment data:

STATE ADMINISTRATIVE DATA FOR WIOA PROGRAMS

Where can I find more information?

The DOL Workforce Report includes a glossary with definitions for common performance measures (common to the core workforce programs) for adults and youth and also some program-specific outcomes. You can find the report here:
https://www.doleta.gov/performance/results/pdf/DOL_Workforce_Rprt_Sep_2015.pdf

State workforce agencies often make annual reports available on their website. These reports also illustrate what data are available and what measures can constructed from the data. Here is an example from the Boston Mayor's Office of Workforce Development:
http://owd.boston.gov/wp-content/uploads/2015/10/2015-Office-of-Workforce-Development-Annual-Report-Booklet_Final.pdf

UNEMPLOYMENT INSURANCE (UI) DATA

What is it?

States operate Unemployment Insurance (UI) programs to provide income during periods of unemployment. States each have their own laws to govern eligibility for UI and how the program is administered, and employers who participate in the UI system report wages of their employees to states (i.e., the overseeing agency or department within the state).

The administration of state UI programs requires quarterly earnings information for all covered employees. This information is referred to as *UI earnings records*.

UNEMPLOYMENT INSURANCE (UI) DATA

What can I measure using these data?

- Whether the youth was employed in a given quarter and quarterly earnings for that quarter
- Length of employment spell, such as the number of consecutive quarters the youth was employed
- Average or total quarterly earnings for that employment spell
- Number of jobs held during the quarter in a given quarter
- Whether an individual was employed in a given quarter

UNEMPLOYMENT INSURANCE (UI) DATA

What should I consider before making a data request?

UI wage data (the most commonly sought-after type of UI data) is reported quarterly. That is, employers report the total amount of wages paid to each of their employees during a specific calendar quarter (i.e., January – March is Q1, April – June is Q2, etc.).

Some employers are not required to report wages and the requirements vary by state. For example, in many cases, self-employed individuals, small businesses and religious organizations are exempt from participating in the UI system and also exempt from reporting wages to the UI agency. Federal and military employment is also excluded.

It is impossible to determine an hourly or weekly wage from the quarterly data.

Not all employment is covered in state UI data, such as employment in a border state or federal employment.

UNEMPLOYMENT INSURANCE (UI) DATA

What should I consider before making a data request?

Although there are some agreements between states to share data in instances where an individual living in one state is employed in another, the agreements are not often binding and the data sharing can be unreliable or unavailable to outside requestors.

There are also some large multi-state employers (i.e. Wal-Mart, McDonalds, etc.) that choose to report all of their wages to a single state regardless of where their employees live.

Employers usually are allowed a 1- to 2- quarter lag time in reporting wages. Most wages earned between January and March (Q1), then, are typically not available until at least July 1.

Wages are reported by employers (identified by an employer ID number), so if an individual has multiple jobs, their wages will be reported in separate transactions. Ideally wages for a particular individual could be matched across jobs using the individual's SSN, though this is not always possible.

UNEMPLOYMENT INSURANCE (UI) DATA

How can I gain access to these data?

As with other types of workforce data, access to UI wage data is controlled by state UI agencies, and there are state laws that govern the circumstances under which data may be shared. Some states do not allow third-party access to UI data.

If data are available and the state is amenable to sharing the data for the purpose of a research study, a **Data Use Agreement** needs to be ratified by all involved parties. Expect to undergo state IRB review.

NATIONAL STUDENT CLEARINGHOUSE (NSC) DATA

What is it?

The National Student Clearinghouse is a nonprofit organization that collects student records on college enrollment and degree attainment from participating colleges and universities nationwide.

More than 3,500 private and public post-secondary institutions report enrollment information to NSC. NSC estimates that these institutions represent 95% of students enrolled in degree-granting institutions.

A complete list of institutions reporting enrollment to NSC can be found on NSC's website.

NATIONAL STUDENT CLEARINGHOUSE (NSC) DATA

What can I measure using these data?

The following can be measures for students attending at least one of NSC's participating colleges:

- Dates of post-secondary enrollment, including whether the youth was previously enrolled at a college
- Enrollment status, whether the youth was enrolled full-time or half-time
- College persistence, such as whether the student enrolled in consecutive terms or whether the student withdrew during the term
- Degree attainment, including the type of degree (e.g., BA, BS)

For examples of outcomes measured using NSC data, see their reports available on their website.

NATIONAL STUDENT CLEARINGHOUSE (NSC) DATA

What should I consider before making a data request?

Program participants must have a signed informed consent that includes specific language about obtaining college records from the NSC.

Detailed course information, such as which courses the student enrolled in or grades, are not available from NSC.

Costs are relatively low and based on the an initial fee of $500 and the number of records submitted for matching:

- $1 per record for the first 1,000;
- $0.60 for 1,000-10,000, and
- $0.36 for over 10,000 records

Once approved, the turnaround time is rather quick. Results are typically returned within a week of submission.

OVERVIEW

HOW TO USE THIS GUIDE

TYPES OF DATA

GENERAL GUIDANCE
- Acquiring the Data
- Understanding the Data

EXAMPLES OF WORKFORCE PROGRAM DATA
- Program Data
- Earnings Data
- Education Data

NATIONAL STUDENT CLEARINGHOUSE (NSC) DATA

How do I gain access to these data?

Requests for access requires the submission of the StudentTracker Agreement to the NSC Research Center. If approved, SSNs and names are then uploaded to the NSC's secure file transfer site for matching.

Contact NSC directly to begin the process or to obtain a copy of the StudentTracker Agreement.

Abt ASSOCIATES

BOLD THINKERS DRIVING REAL-WORLD IMPACT