



TANF DATA
COLLABORATIVE

FEBRUARY 2024
OPRE REPORT
2024-005

Richard Hendra
Johanna Walter
Audrey Yu

Transforming Administrative Data into a Resource for Evidence Building

Government agencies collect vast amounts of administrative data in their day-to-day activities, primarily for program operations. But the information is less often used as a research tool or fully harnessed for its evidence-building potential. This brief is the fourth in a series of publications from MDRC about the [Temporary Assistance for Needy Families \(TANF\) Data Innovation \(TDI\) project](#), initiated by the Administration for Children and Families (ACF) in 2017.¹ It describes TDI's efforts to transform federal TANF and employment data into an integrated resource for evidence-informed program management and policymaking. Such efforts are rarely an easy lift. Creating the data resource required deliberate federal investments and the resolution of many technical, legal, and security issues by the TDI team, which was led by MDRC in collaboration with Actionable Intelligence for Social Policy, Chapin Hall at the University of Chicago, and the Coleridge Initiative. This brief describes how the team tackled this challenging project. Although the work was focused on the TANF program, the lessons learned may be useful for members of federal and state agencies, researchers, and advocates of evidence-informed policy making interested in unlocking the potential of administrative data.

Transforming administrative data into a research-ready format is challenging for a number of reasons. It often requires extensive data processing to overcome issues such as a lack of standardization, poor documentation, and storage in non-research-oriented formats.² Sometimes, key learning questions require merging data across multiple systems, which may be labor intensive. In the case of TANF cash assistance data, for example, long-term employment outcomes are seldom available in the same data system as information on TANF benefits receipt, despite self-sufficiency being the goal of the TANF program.³ The conversion process demands considerable effort and time, often involving many hours to organize, understand, and document the data. This work is not easily automated and is sometimes discounted or not fully understood by designers of data initiatives. Often, the privacy of the data subjects (in the TDI project that meant TANF recipients and their families) is a critical concern requiring the establishment of meticulous safeguards.⁴ Agencies may simply lack the necessary resources to overcome these issues. Outdated technology and organizational structures can also hinder the process.⁵

To tackle capacity issues at the federal level, several MDRC staff members on the TDI team were embedded within ACF, which enabled them to identify gaps and support improvement of data management and data processing within the agency. Data governance challenges were managed by forming a series of data-sharing agreements that adhered to strict data security protocols. To enhance data quality, documentation, and standardization, the TDI team and ACF collaborated closely on data cleaning, transformation, and the documentation necessary for research use of the data. Replicable and scalable data files and analysis tools were also developed to aid the research process. ACF's National Directory of New Hires (NDNH) employment data were linked with TANF data to support the measurement of employment outcomes. To manage legacy data systems and large NDNH files, extensive testing and file segmentation were undertaken. Lastly, to ensure secure data access, the data were prepared and accessed within the Coleridge Initiative's Administrative Data Research Facility (ADRF), which provides stringent data security and access tools.

Despite significant progress, inherent limitations remain regarding the use of TANF data for research because of reporting flexibilities allowed by law. The most prominent limitation for research use is that states and territories (referred to in this brief as "states") are allowed to report a sample of cases instead of their entire caseload.⁶ This precludes longitudinal analysis of benefits receipt in "sample" states and limits the data's full potential for analysis and evidence building.

Components of the TANF Data Innovation Project

The goal of the TDI project, sponsored by the Office of Family Assistance (OFA) and the Office of Planning, Research, and Evaluation (OPRE) within ACF, was to enhance the data analytics capacity of TANF agency staff members at the federal and state levels using approaches such as training, technical assistance, coaching, and the development and enhancement of the federal data infrastructure. The project consisted of three main components:

- **The TANF Data Collaborative (TDC)** supported the efforts of TANF agencies to routinely use administrative data to inform policy and practice through targeted training and technical assistance (TA). The component also included the TDC Pilot Initiative—an intensive TA program for staff members at eight state TANF agencies.
- The **Federal Capacity-Building (FedCap)** component provided ACF with technical assistance and staffing support to augment federal staff members' ability to improve the quality, use, and analysis of data within OFA's TANF Data Division.
- The ongoing **TANF Employment Project (TEP)** focuses on refining TANF data infrastructure to support evidence building. This includes administrative data transformation, documentation, and analysis to develop longitudinal, individual-level TANF data files suitable for analytics use. The project also integrates data on TANF benefits receipt with longitudinal earnings data from the NDNH. Additionally, ACF, through TEP, contracted for a secure platform to provide remote access to data files and established the necessary governance protocols to enable researcher access.

Improving Data Representativeness and Quality by Building Federal Capacity

FedCap was a strategic collaboration between ACF and an embedded group of TDI team members from MDRC. The aim was to strengthen ACF's ability to enhance the gathering, management, and analysis of TANF data from state agencies. Key goals included easing the TANF data reporting process, improving the quality of the data received from state TANF agencies, and aiding OFA in effectively using federally reported TANF data. The first stage of FedCap activities focused on improving TANF data quality and representativeness. This included creating dashboards and reports to identify quality issues, developing technical resources to support quality control processes for states, and enhancing federal data process documentation. The team's findings also contributed to the development of an upgraded TANF data reporting system.

As the project advanced, FedCap activities shifted from TDI staff members working with ACF onsite to working remotely. Their efforts included developing automated fact sheets for OFA regional administrators, states, and territories; crafting a memo to guide TANF researchers on possible data uses; and conducting a comprehensive review of TANF sampling procedures, which led to a guidance document for states that are currently using sampling or are considering using it in the future.

Enabling Secure Access to TANF and Employment Data and Analysis Templates

While the FedCap initiative laid the groundwork for using administrative data for evidence building, TEP has moved these efforts forward. The TEP initiative, which runs through mid-2024, aims to advance the analysis of TANF data in order to understand the trajectories and employment outcomes of TANF program participants. The TDI team is developing two key data resources: a national, longitudinal file of TANF benefits receipt that covers fiscal year 2009 onward, and integrated TANF and NDNH data for two historical cohorts

and one ongoing cohort. The TANF/NDNH file enables analysis of program participants' employment outcomes, including comparisons of current TANF recipients' outcomes with those of past participants.

To facilitate data access while protecting the privacy of program participants, these data are securely stored in the Coleridge Initiative's ADRF. Authorized users can access data through a remote interface, but data cannot be transferred out of ADRF. Only aggregate information can be released after a strict review process.

In the future, these data resources might be accessible to other researchers, providing insights to enhance TANF programs. To make the data more usable, the TDI team is providing sample code for data quality checks, processing, and data analysis, along with extensive documentation. This will help users quickly and effectively start working with the data. Eventually, TEP might enable authorized users to evaluate long-term employment outcomes of TANF recipients.

TEP Data Assets

As mentioned, TEP is developing two major data resources: TANF longitudinal data and integrated TANF and employment data.

The ***TANF longitudinal data*** resource aims to facilitate research on TANF usage patterns, including how economic factors and state policies are related to benefits receipt. This has become critical in light of recent economic volatility and the COVID-19 pandemic, which may have influenced the demographics and needs of TANF enrollees (see Box 1). In addition, as noted in qualitative interviews with national TANF experts, there has arguably been a lack of research focused on TANF in recent years, which means the area is ripe for evidence building.⁷ The TEP TANF data are organized into different research-ready formats and are heavily documented. Because every state has a unique TANF program, standardizing TANF records requires a lot of effort. In addition, the preparation of TANF data requires careful decisions to define the unit of analysis, such as what constitutes a household. TEP (and work conducted by Chapin Hall before it) has made significant progress in developing standardized definitions, making these data more usable for evidence building.

The ***Integrated TANF and employment data*** resource combines TANF data with employment data from the NDNH. The establishment of this dataset required a collaborative effort between two federal entities—ACF's OFA and the Office of Child Support Services (OCSS). This comprehensive dataset can support routine performance measurement, research, and the analysis of potential enhancements to performance measures for the TANF program. It can also offer insights into the employment retention and progression patterns of TANF recipients. Access to and acceptable use of these data are strictly regulated by memos of understanding between OFA and OCSS. (See Box 2 for more about NDNH data.)

Box 1. What is TANF Administrative Data?

Temporary Assistance for Needy Families (TANF) is a federal cash assistance program that was enacted in 1996 to provide short-term financial aid to families with low incomes. The program replaced the Aid to Families with Dependent Children program.* TANF functions through block grants to states, tribes, and territories. TANF agencies must report monthly data and file quarterly reports to the Administration for Children and Families within the Department of Health and Human Services, providing details such as the TANF benefit amount received by the participant, services provided, and demographic and work training activity information. This data is used to calculate the work participation rate, the main metric for state accountability. TANF legislation grants states flexibility to report data for their entire caseload or a subset using approved sampling methods.†

NOTES: *Office of Family Assistance, "About TANF" (Administration for Children and Families, U.S. Department of Health and Human Services, 2022). Website: <https://www.acf.hhs.gov/ofa/programs/tanf/about>.

†Office of Family Assistance, "TANF-ACF-PI-2020-04 (Form ACF-199, TANF Data Report, and Form ACF-209, SSP-MOE Data Report, Approved by the Office of Management and Budget through 11/30/2023, OMB Control Number 0970-0338" (Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services, 2022). Website: <https://www.acf.hhs.gov/ofa/policy-guidance/tanf-acf-pi-2020-04>.

Enhancing Research Capabilities through TEP's Data Preparation and Tool Development

Beyond navigating the complexities of data governance, the TEP initiative also encompasses understanding, standardizing, and reshaping a vast array of TANF data into a comprehensible and usable format. The project includes detailed data documentation and proof-of-concept analyses, with the aim of creating a robust data infrastructure to facilitate policy-relevant research and a clearer understanding of the intricacies of the TANF program.

- **DATA PREPARATION AND DOCUMENTATION.** Building the TANF data infrastructure involves more than just data governance; it requires interpreting, standardizing, and restructuring complex, state-specific data into a research-friendly format. This includes the development of "spells" or well-defined periods of TANF benefits receipt, and cleaning irregular data such as separate state program (SSP) counts. The TDI team has also prepared detailed data dictionaries and associated documentation for a reliable and comprehensive understanding of the data.
- **INITIATING TEP PROOF-OF-CONCEPT ANALYSES.** The TDI team aims to illustrate the potential of the TANF and employment data through a series of proof-of-concept analyses. The goal is to provide federal staff and researchers with practical examples and insights about the data, its challenges, and opportunities.

Box 2. What's in the National Directory of New Hires?

The NDNH is a valuable administrative database that was established by the federal Personal Responsibility and Work Opportunity Reconciliation Act of 1996 to support state efforts in child support enforcement. The NDNH compiles quarterly employment and earnings information including data covered by state Unemployment Insurance systems and federal entities. The NDNH comprises three types of data:

- **THE NEW HIRE FILE** contains information on newly employed individuals as reported by employers to their respective State Directory of New Hires.
- **THE QUARTERLY WAGE FILE** provides individual earnings data derived from state workforce and federal agency records.
- **THE UNEMPLOYMENT INSURANCE FILE** offers information on individuals who have either received or applied for unemployment benefits.

SOURCE: Office of Child Support Services, "A Guide to the National Directory of New Hires" (Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services, 2023). Website: <https://www.acf.hhs.gov/css/training-technical-assistance/guide-national-directory-new-hires>

NOTE: Studies indicate that independent contractor (Internal Revenue Service 1099) roles, which are included in NDNH reporting at states' discretion, account for 6 to 14 percent of employment. This percentage may fluctuate over time based on measurement methods (see Lawrence F. Katz and Alan B. Krueger, "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015," NBER Working Paper No. 22667 (Cambridge, MA: National Bureau of Economic Research, 2016); U.S. Bureau of Labor Statistics, "Contingent and Alternative Employment Arrangements – May 2017" (Washington, DC: U.S. Bureau of Labor Statistics, U.S. Department of Labor, 2018); Katharine G. Abraham, John C. Haltiwanger, Kristin Sandusky, and James R. Spletzer, "Measuring the Gig Economy: Current Knowledge and Open Issues," NBER Working Paper No. 24950 (Cambridge, MA: National Bureau of Economic Research, U.S. Bureau of Labor Statistics, 2018).

- **BUILDING-BLOCK ANALYSES OF TANF LONGITUDINAL DATA.** The TDI team is developing foundational materials for TANF researchers. The focus is on establishing definitions of measures and concepts vital to TANF data analysis. The team aims to simplify complex research questions and to guide new researchers in how to use TANF data effectively.
- **COHORT-BASED ANALYSES OF INTEGRATED TANF AND EMPLOYMENT DATA.** The TDI team will also prepare materials for the integrated TANF and employment data resource, providing documentation, quality checks, and data processing. The team aims to conduct example analyses using three distinctive TANF recipient cohorts, sharing all programming code to aid future researchers.
- **STATE EMPLOYMENT REPORTS:** A key planned data product includes state-level quarterly and annual employment outcome reports covering TANF recipients and leavers.

These reports will aid states in tracking TANF participants' employment progress and will be accessible without special access to the ADRF.

The TDI team will also explore a variety of research questions related to TANF employment retention and advancement.

Conclusion

Creating data assets for expanded TANF data analysis and evidence building required the TDI team to overcome challenges in multiple areas, including organizational capacity, data representativeness and quality, data availability, data security, and legal requirements.

Table 1 presents a review of the challenges of transforming TANF data into an accessible resource, how these challenges manifested in the context of the TDI project, and how the team sought to overcome them.

Table 1. Challenges and Solutions for Building TANF Data Infrastructure

CATEGORY	CHALLENGE	SOLUTION
Organizational capacity	Staff time constraints hindered new projects. Staff members lacked the time, technical skills, and access to technology to transform data for evidence building.	TDI provided on-site contractors who identified obstacles and offered complete support for project tasks. They enhanced data management capacity and helped in building scalable data files and tools for problem-solving.
	Existing data infrastructure couldn't support new workstreams.	TDI collaborated with federal staff members to test and iterate data flows, enabling workarounds like dividing larger files into smaller ones.
Data quality	Some states use nonrepresentative TANF data samples, limiting the data's usefulness in longitudinal research.	TDI conducted response analyses and provided recommendations for improving sampling processes and identifying research questions that were answerable with existing data.
	There were data quality issues from disparate reporting methods, missing or duplicated data, outliers, and identifier problems.	TDI developed standardized quality control programs to tackle these concerns, including establishing new household identifiers.
	Making state-specific data useful for national research could be difficult due to variation in TANF benefits and services. ^a	TDI is creating standardized measures and potential approaches to multistate analyses.

(continued)

Table 1 (continued)

CATEGORY	CHALLENGE	SOLUTION
Data availability	TANF data doesn't track participants' long-term employment.	TDI linked TANF data to the Administration for Children and Family's (ACF) National Directory of New Hires employment data, which saved individual states from having to make data agreements—simplifying access for states to aggregate statistics and allowing standardized employment measures to be developed.
Data security and legal frameworks	Analysts had to balance the usefulness of sensitive TANF and employment data with privacy and security requirements.	Data were linked using anonymized keys, with sensitive data elements masked, and stored and accessed in a secure cloud-based platform, with rigorous security protocols approved by the office of the chief information officer in ACF.
Data documentation and onboarding	Poor documentation and a steep learning curve made federal administrative data challenging for new analysts to use.	TDI produced detailed documentation and examples for researchers, published reports on the project's work, and plans to create state summary reports to disseminate research insights.

SOURCE: ^aMeasuring employment outcomes in light of these differences is discussed in Office of Planning, Research and Evaluation, "Measuring Employment Outcomes in TANF" (Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services, 2018). Website: <https://www.acf.hhs.gov/opre/report/measuring-employment-outcomes-tanf>

The TDI project showcases a successful approach to using administrative data for research and evidence building. The collaborative, interdisciplinary TDI team made significant progress toward developing a strong infrastructure with comprehensive documentation and programming tools. The project required strategic planning, careful resource allocation, and cooperation at various administrative levels. Ensuring responsible data management and fostering deep and aligned collaborations are key to converting administrative data into an evidence-building resource.

Notes and References

- 1 Office of Planning, Research and Evaluation, "Temporary Assistance for Needy Families (TANF) Data Innovation Project" (Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services, 2023). Website: <https://www.acf.hhs.gov/opre/project/tanf-data-innovation-project-2017-2024>.
- 2 Robert M. Groves and George J. Schoeffel, "Use of Administrative Records in Evidence-Based Policymaking," *The Annals of the American Academy of Political and Social Science* 678, 1 (2018): 71-80; Scott W. Allard, Emily R. Wiegand, Colleen Schlecht, A. Rupa Datta, Robert M. Goerge, and Elizabeth Weigensberg, "State Agencies' Use of Administrative Data for Improved Practice: Needs, Challenges, and Opportunities," *Public Administration Review* 78, 2 (2018): 240-250.
- 3 Robert M. Goerge, Emily R. Wiegand, and Leah Gjertson, "Unpacking Data Use In State TANF Agencies," OPRE Report 2021-94 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021); Edith Yang, Sharon Zanti, T.C. Burnett, Richard Hendra, Dennis Culhane, Zarni Htet, Della Jenkins, Camille Prael-Dumas, and Electra Small, "Expanding TANF Program Insights: A Toolkit for State and Local Agencies on How to Access, Link, and Analyze Unemployment Insurance Wage Data, OPRE Report 2022-226 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2022).
- 4 U.S. Commission on Evidence-Based Policy Making, "Barriers to Using Administrative Data for Evidence-Building" (2016). Website: <https://www.acf.hhs.gov/opre/project/commission-evidence-based-policymaking-cep>; Goerge, Wiegand, and Gjertson (2021); Edith Yang et al. (2022).
- 5 Nick Hart and Carmody Kody, "Barriers to Using Government Data: Extended Analysis of the U.S. Commission on Evidence-Based Policymaking's Survey of Federal Agencies and Offices" (Washington, DC: Bipartisan Policy Center, 2018); Robert M. Goerge, "Barriers to Accessing State Data and Approaches to Addressing Them," *The Annals of the American Academy of Political and Social Science* 675, 1 (2018): 122-137; Melissa Wavelet, Stephanie Rubino, Caroline Morris, and Maryah Garner, "Increasing Data Analytics Capacity in State TANF Agencies: The TANF Data Collaborative Approach," OPRE Report 2022-124 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2022).
- 6 Administration for Children and Families, *TANF Sampling and Statistical Methods Manual* (Washington, DC: U.S. Department of Health and Human Services, 2007); Office of Family Assistance, "TANF-ACF-PI-2020-04, Form ACF-199, TANF Data Report, and Form ACF-209, SSP-MOE Data Report, Approved by the Office of Management and Budget through 11/30/2023, OMB Control Number 0970-0338," (Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services, 2020). Website: <https://www.acf.hhs.gov/ofa/policy-guidance/tanf-acf-pi-2020-04>; Brett Brown and Robin Gherter (eds.), *A Compendium of Administrative and Survey Data Resources in the Administration for Children and Families*, OPRE Report 2020-10 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2020).
- 7 One of the findings from interviews with stakeholders was that TANF-related research peaked in the period just after the 1996 welfare law but has been on the wane in recent years. For more details see Goerge, Wiegand, and Gjertson (2021). The brief summarizes results from a 2019 needs assessment of the capacity of TANF programs in 54 U.S. states and territories to analyze data used for the purposes of program improvement, monitoring, and evidence building. It highlights areas of strength and success in how these agencies use data, as well as areas for growth. It also includes suggested strategies that may improve data use by TANF agencies.

Acknowledgments

We would like to express our most sincere appreciation to the individuals whose insights contributed to this brief. Special thanks to Nicole Deterding, Kathleen Moore, Alexandra Pennington, Siri Warkentien, Bethanne Barnes, and Lauren Frohlich of the Administration for Children and Families in the U.S. Department of Health and Human Services for informing, reviewing, and supporting this brief. We would like to thank our partners on the TANF Data Innovation Team— Actionable Intelligence for Social Policy, the Coleridge Initiative, and Chapin Hall. Particular thanks are due to Erika Lundquist and Melissa Wavelet of MDRC, Maryah Garner of the Coleridge Initiative, and Emily Weigand at Chapin Hall for multiple reviews and edits. Thanks to Chris Homa and Sufiyan Syed for coordinating, Jill Kirschenbaum for editing the brief, and Carolyn Thomas for preparing the brief for publication.

OPRE REPORT 2024-005

FEBRUARY 2024

AUTHORS: Richard Hendra, Johanna Walter, and Audrey Yu

SUBMITTED TO: Rafael J. Elizalde, Sr., Project Officer, Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services

CONTRACT NUMBER: HHSP233201500059I

This report is in the public domain. Permission to reproduce is not necessary.

SUGGESTED CITATION: Hendra, Richard, Johanna Walter, Audrey Yu. 2024. "Transforming Administrative Data into a Resource for Evidence Building." OPRE Report No. 2024-005. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

DISCLAIMER: The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research, and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services. This report and other reports sponsored by the Office of Planning, Research, and Evaluation are available at www.acf.hhs.gov/opre.

Connect with OPRE



The TANF Data Innovation project is funded by the Office of Family Assistance and the Office of Planning, Research, and Evaluation within the Administration for Children and Families. MDRC and subcontractors Actionable Insights for Social Policy at the University of Pennsylvania, Chapin Hall at the University of Chicago, and Coleridge Initiative are conducting the Promoting and Supporting Innovation in TANF Data project for the Department of Health and Human Services (HHS) under a contract with the Administration for Children and Families, funded by HHS under a competitive award, Contract No. HHSP233201500059I. The project officer is Rafael J. Elizalde.

Dissemination of MDRC publications is supported by the following organizations and individuals that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Arnold Ventures, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, and Sandler Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

NEW YORK

200 Vesey Street
23rd Flr.
New York, NY 10281
Tel: 212 532 3200

OAKLAND

475 14th Street
Suite 750
Oakland, CA 94612
Tel: 510 663 6372

WASHINGTON, DC

750 17th Street, NW
Suite 501
Washington, DC 20006

LOS ANGELES

11965 Venice Boulevard,
Suite 402
Los Angeles, CA 90066

