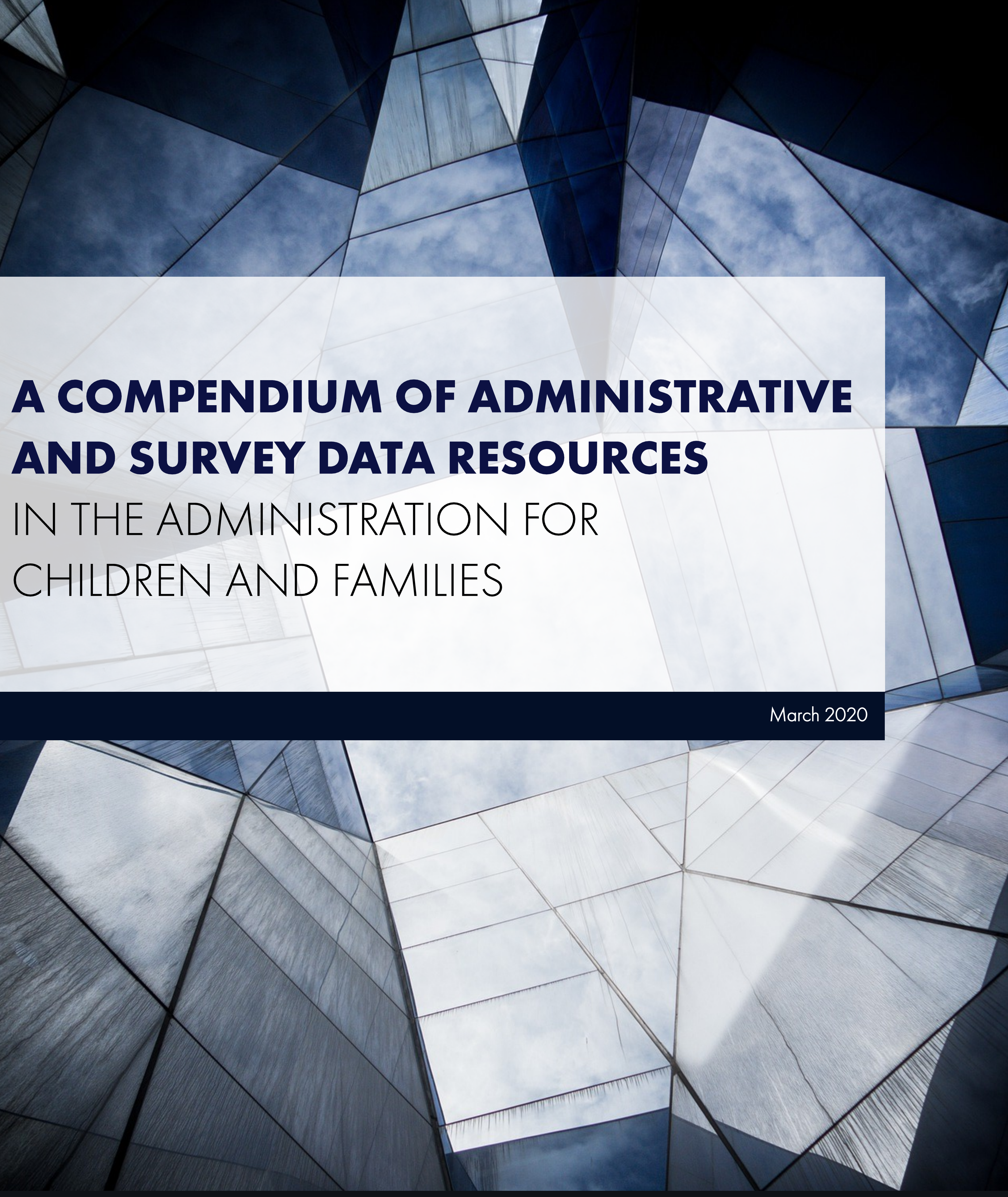

A COMPENDIUM OF ADMINISTRATIVE AND SURVEY DATA RESOURCES
IN THE ADMINISTRATION FOR CHILDREN AND FAMILIES

March 2020

A COMPENDIUM OF ADMINISTRATIVE AND SURVEY DATA RESOURCES IN THE ADMINISTRATION FOR CHILDREN AND FAMILIES

OPRE Report 2020-10

March 2020

A joint product of
The Division of Data and Improvement,
Office of Planning, Research, and Evaluation
and
The Office of the Assistant Secretary for Planning and Evaluation

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Brown, Brett, and Robin Ghertner (eds) (2020). *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.

Sign-up for the
ACF OPRE News
E-Newsletter

Like OPRE on Facebook
facebook.com/OPRE.ACF

Follow OPRE on Twitter
@OPRE_ACF

TABLE OF CONTENTS

INTRODUCTION

There has been a sustained focus on how to make better use of federal administrative data to conduct research and evaluation of government-funded programs in order to improve outcomes and efficiency for the people we serve. Research and evaluation are not separate from the business of program operations and strategy, but must go hand-in-hand with other efforts to ensure programs and policies meet their goals and adapt to new challenges. Making full use of rich data that has already been collected can greatly increase our knowledge about how to improve performance and effectiveness without additional burden on states, communities, or other grantees or requiring significant new resources.

The Foundations for Evidence-Based Policymaking Act of 2018 requires the development of a federal data catalogue and includes other provisions requiring agencies to make information about data they collect available to the public. The Commission on Evidence-Based Policymaking issued a report in 2017 emphasizing the use of administrative data sources, and laid out several recommendations for how such data can be better used.[1] The Office of Management and Budget (OMB) has issued several memoranda on administrative and survey data for evidence-building activities.[2]

The Administration for Children and Families (ACF) collects a wide range of data from grantees and other stakeholders. Though these data are used primarily to manage programs, they provide rich opportunities to conduct further analysis and research. ACF has long relied on these data sources to better understand its program operations and outcomes, and in some cases, conduct evaluations. These efforts have already had some success, and more can be done to build on that success. In addition, ACF sponsors various survey data collections in order to understand specific programs and populations, resources which are often enhanced by linking to administrative records from both within and beyond ACF.

---

[1] Commission on Evidence-Based Policy (CEP). “The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking.” Final Report. Washington, DC: CEP, September 2017. Available at https://www.cep.gov/cep-final-report.html (accessed March 1, 2019).

[2] Office of Management and Budget. “Barriers to Using Administrative Data for Evidence-Building.” White paper prepared by the Office of Management and Budget for the Commission on Evidence-Based Policymaking. Washington, DC: Office of Management and Budget, July 15, 2016.

Office of Management and Budget. “Guidance for Providing and Using Administrative Data for Statistical Purposes.” OMB M-14-06. Washington, DC: Office of Management and Budget, 2014.

This *Compendium of Administrative and Survey Data Sources in the Administration for Children and Families* documents the data collected by ACF that is or could be used for evidence-building purposes. Among other uses, it can serve as a reference point for efforts to maximize the use of data for analysis, program improvement, evidence-building, and other purposes. Further analysis of the characteristics of the data sources can identify challenges and opportunities for data sharing, linking, and evidence-building. The compendium can also inform ACF evaluation activities helping to identify which data sources may be useful for which evaluation questions. Of note, many of the datasets are not available to individuals outside of the federal government due to statutory and regulatory restrictions. While each data source has its own authorizing statute or regulation, many also fall under privacy restrictions established under broader federal statutes, such as the Health Insurance Portability and Accountability Act (HIPAA), the Family Educational Rights and Privacy Act (FERPA), and the Privacy Act of 1974.

To compile the information in the Compendium, we consulted subject matter experts – program and data experts within the U.S. Department of Health and Human Services (HHS) with extensive knowledge of program administrative and survey data related to ACF programs. First, we held discussions with the experts to identify the data sources to be included. To be included, a dataset had to: 1) include person or case-level data; 2) relate directly to an ACF program; 3) have the potential to be used for analytical or evaluation purposes to improve program operations; and 4) be an administrative database with national scope or a nationally representative survey. Second, the project team defined the meta-data elements to be collected about each dataset, and developed a standardized data collection instrument (DCI) for these elements. Next, the team worked with key informants within the Office of Planning, Research, and Evaluation (OPRE) and the Office of the Assistant Secretary for Planning and Evaluation (ASPE) to complete the DCI for each data source. Finally, experts from program offices provided quality review of the DCI for the data sources under their purview.

Each entry in the Compendium includes: an overview of the data source; data ownership and funding source; basic content (topical areas covered); major publications, websites, and documentation; available datasets (public and restricted); data quality; statutory and regulatory restrictions on access and use; capacity to link with other data sources; and examples of prior research using linked data. General guidance on appropriate content for each category was provided to the authors, along with examples, though they were given latitude to include what they thought was most important. As a result, the entries are rich in content, but the content may vary significantly across them. Where appropriate, existing descriptions from available federal sources were used as the basis for the descriptions.

Two main limitations to the Compendium that should be kept in mind. First, while it captures the majority of primary administrative and survey datasets related to ACF programs, it does not include every dataset that could meet our criteria at some point.

For example, there are existing data sources used by ACF programs that do not meet our third criterion, but that in the future may serve that purpose. Second, the information in this Compendium is likely to change over time. ACF programs are constantly evolving, and as they do new data sources may be developed while existing ones may become defunct. Content of data sources is also likely to change, as well as how the data sources are used. This all means that the Compendium is a snapshot at a specific time.

DATA ELEMENTS REPORTED

Each entry in the Compendium includes several elements about the data source in question. These elements include:

Website: The primary website to learn more about the data.

Funding Agency and Data Ownership: The agency that funds and legally owns the data.

Overview: A short description of the data source and programs covered by the data source.

Periodicity: The frequency with which the data are collected and made available. For example, some data sources are collected annually from grantee reports, while others are continuously updated.

Data Content and Structure: An overview of the main data elements included in the dataset.

Major Publications: Examples of significant publications using the data. These publications provide insight into how the data can be used for research or other purposes.

Available Datasets: Information on what data products are available, including those available to the public with no restrictions, as well as restricted-use datasets requiring specific permission to access.

Statutory and Regulatory Restrictions: Restrictions on how the data can be used, provided in statute and agency regulation.

Feasibility to Link to Other Data: An assessment of how and whether the data could be linked to other data sources.

Data Quality: Information on the quality of the data collected, including validity and reliability. Generally, formal data quality reviews or assessments have not been conducted on many of the datasets. Information on data quality for most datasets come from subjective opinions of federal experts who manage the data or use the data frequently, along with objective measures such as non-response rates and population coverage.

Program Scope and Budget: Information on the general size of the programs covered by the data, including recent budgetary allocations or annual appropriations.

LIST OF ACRONYMS

ACA	Affordable Care Act
ACF	Administration for Children and Families
ACS	American Community Survey
ACYF	Administration for Children, Youth, and Families
ADVHOCaT	Accomplishments of the Domestic Violence Hotline, Online Connections, and Text
AFCARS	Adoption and Foster Care Analysis and Reporting System
AIR	American Institutes for Research
APPRISE	Applied Public Policy Research Institute for Study and Evaluation
ASPE	Office of the Assistant Secretary for Planning and Evaluation
ASR	Annual Survey of Refugees
ATO	Authorization to Operate
Baby FACES	Early Head Start Family and Child Experiences Survey
BCP	Basic Center Program
BRS	Bayley Behavior Rating Scale
CAPTA	Child Abuse Prevention and Treatment Act
CARRA	Center for Administrative Records Research and Applications
CCDBG	Child Care and Development Block Grant Act
CCDF	Child Care and Development Fund
CDC	Centers for Disease Control and Prevention
CFCIP	John Chafee Foster Care Independence Program
CFSR	Child and Family Services Reviews

CIPSEA	Confidential Information Protection and Statistical Efficiency Act
CoC	Continuum of Care
CPS	Child Protective Services
CSPED	Child Support Noncustodial Parent Employment Demonstration
DCI	data collection instrument
DOB	date of birth
ECE/SA	early care and education/school-age
ECI	Early Communication Indicator
EIA	U.S. Energy Information Administration
ESS	Energy Supplier Survey
ETJD	Enhanced Transitional Jobs Demonstration
ETV	Education and Training Voucher Program
FACES	Head Start Family and Child Experiences Study
FCR	Federal Case Registry
FEIN	Federal Employer Identification Number
FERPA	Family Educational Rights and Privacy Act
FPLS	Federal Parent Locator Service
FVPSA	Family Violence Prevention and Services Act
FY	fiscal year
FYSB	Family Youth Services Bureau
HHS	U.S. Department of Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act
HOME	Home Observation for Measurement of the Environment
HOVRS-A	Home Visit Rating Scales-Adapted

HPOG	Health Profession Opportunity Grants
HUD	Housing and Urban Development
ICPSR	Inter-university Consortium for Political and Social Research
IDEA	Individuals with Disabilities Education Act
IDS	integrated data systems
IRB	Institutional Review Board
LEHD	Longitudinal Employer-Household Dynamics
LIHEAP	Low Income Home Energy Assistance Program
LIR	loveisrespect
MGH	Maternity Group Home
MiHOPE	Maternal and Infant Home Visiting Program Evaluation
MSFIDM	Multistate Financial Institution Data Match
NCANDS	National Child Abuse and Neglect Database
NCHS	National Center for Health Statistics
NDACAN	National Data Archive on Child Abuse and Neglect
NDNH	National Directory of New Hires
NEWWS	National Evaluation of Welfare-to-Work Strategies
NICHD	National Institute of Child Health and Human Development
NIH	National Institutes of Health
NIS	National Incidence Study
NLNPCG	No Longer Nonparental Caregiver
NPRM	Notice of Proposed Rulemaking
NSCAW	National Survey of Child and Adolescent Well-Being
NSCH	National Survey of Children’s Health

NSECE	National Survey of Early Care and Education
NYTD	National Youth in Transition Database
OCC	Office of Child Care
OCCIS	Office of Child Care Information System
OCSE	Office of Child Support Enforcement
ODARE	Office of Data, Analysis, Research, and Evaluation
OFA	Office of Family Assistance
OMB	Office of Management and Budget
OPRE	Office of Planning, Research, and Evaluation
ORR	Office of Refugee Resettlement
PACE	Pathways for Advancing Careers and Education
PACT	Parents and Children Together Evaluation
PII	personally identifiable information
QW	quarterly wage file in the National Directory of New Hires
R&P	reception and placement
RADS	Refugee Arrivals Data System
RECS	Residential Energy Consumption Survey
RHY-HMIS	Runaway and Homeless Youth – Homeless Management Information System
RPG	regional partnership grant
SAFE-T	Summative ADEPT Formal Evaluation of Classroom-Based Teachers
SDS	Sentinel Definitions Survey
SLAITS	State and Local Area Integrated Telephone Survey
SNAP	Supplemental Nutrition Assistance Program

SOR	system of record
SPM	Structure and Practices Mail Survey
SPS	Screening Policies Study
SSA	Social Security Administration
SSN	Social Security number
SSP	separate state programs
STED	Subsidized and Transitional Employment Demonstration
SWA	state workforce agency
TANF	Temporary Assistance for Needy families
TLP	Transitional Living Program
UI	unemployment insurance
USC	United States Code
WIA	Workforce Investment Act
WRAPS	Worldwide Refugee Admissions Processing System
WRMA	Walter R McDonald & Associates, Inc.

ACF ADMINISTRATIVE DATA SOURCES

Adoption and Foster Care Analysis and Reporting System (AFCARS)

Website

https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/afcars

Funding Agency and Data Ownership

Children's Bureau, Administration on Children, Youth, and Families (ACYF), ACF/HHS

Data Gathering Agency/Contractor

Data are gathered and reported by the states. Data are housed on a National Institutes of Health (NIH) server, downloaded and processed by the Children's Bureau, and analyzed by ACYF's Office of Data, Analysis, Research and Evaluation (ODARE).

Overview

AFCARS collects case-level information from state and tribal Title IV-E agencies on all children in foster care and those who have been adopted with Title IV-E agency involvement. AFCARS was designed to address policy development and program management issues at both state and federal levels. The data are also useful for researchers interested in analyzing aspects of the United States' foster care and adoption programs.

(Sources: https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/afcars and https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/AFCARSGuide2000toPresent.pdf.)

Periodicity

Title IV-E agencies are required to submit AFCARS data semi-annually to the Children's Bureau. The AFCARS report periods are October 1 through March 31 and April 1 through September 30. Data for each report period are due no later than May 15 and November 14, respectively.

Data Content and Structure

Topics reported in the AFCARS foster care file include:

- Demographic information on the foster child as well as the foster parents
- FIPS code (number identified for the county that has responsibility for the case)
- The number of removal episodes a child experienced

- The number of placements in the current removal episode
- The current placement setting type
- Circumstances associated with removal from the home
- Disability information
- Dates of removal and dates of discharge
- Reasons for discharge from foster care
- Whether the child is receiving Title IV-E foster care payments, adoption assistance, child support, Medicaid, or Social Security Act benefits

Additional data available in the adoption file include:

- Whether the child had special needs and which types
- Year of birth for biological parents
- Demographic information about adoptive parents
- Adoptive parents' relationship to child
- Date of adoption finalization

For a complete listing of elements, definitions, policy clarifications, and other supporting information for each element, see AFCARS Assessment Review Guide, Appendix C: https://www.acf.hhs.gov/sites/default/files/cb/afcars_assessment_review_guide.pdf

The National Data Archive on Child Abuse and Neglect (NDACAN), funded by the Children's Bureau, has resource material including a code book for those agencies receiving data files for research purposes. https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/AFCARSAdoptionCodebook.pdf

Major Publications

Summary statistics related to numbers of children in foster care and who are adopted are available on the Children's Bureau's website at https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/afcars

In addition, the Children's Bureau prepares the annual Child Welfare Outcomes Reports to Congress and maintains a data site at https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/cwo. This report and data site includes data from AFCARS, reported at the state level.

Available Datasets

Public-use aggregate:

Aggregate statistics are available on the Children's Bureau website: https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/afcars.

Tabular data can be made available upon request from NDACAN at http://www.ndacan.acf.hhs.gov.

Public-use micro:

None (microdata files are available to researchers only, see below).

Restricted-use micro:

NDACAN makes AFCARS data available upon request. AFCARS data are restricted access files, as only individuals with a research-related affiliation with an institution may request data. Requestors must provide detailed contact information and sign a Terms of Use agreement that stipulates that the investigator must have a research-related affiliation with an institution or be granted an exception from NDACAN. Undergraduate students may serve as investigators if a faculty advisor co-signs the agreement. They must assure confidentiality, prevent unauthorized disclosure of confidential information, never make intentional identification of a research subject, and only use the data solely for research or statistical purposes. In addition, NDACAN staff mask the data prior to distribution. To mitigate the risk of disclosure, dates of birth are rounded and counties with fewer than 1,000 children are recoded. All unique identifiers are encrypted and no personally identifiable information (PII) is included.

A codebook with more detail on the masking procedure can be found at https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/AFCARSGuide2000toPresent.pdf

In-house micro:

In-house datasets contain all state and county identifiers, as well as original values for measures that are rounded, suppressed, or perturbed in the datasets distributed by NDACAN. Staff may use these datasets to respond to specific inquiries from federal agencies or Congress. In addition, they have occasionally partnered with outside researchers for joint research efforts that involve use of these datasets. Those partnerships are governed by specifically negotiated data use agreements.

Statutory and Regulatory Restrictions

AFCARS is a statutorily mandated collection of the demographics and experiences of children in foster care. AFCARS regulations were initially promulgated in 1993 (under 45 CFR 1355.40 and the appendices of 1355) and were first updated through a final rule published in December 2016 (81 FR 90524). The December 2016 final rule substantially increased the number of data points that must be reported. ACF published a Notice of Proposed Rulemaking (NPRM) on April 19, 2019, to streamline the AFCARS data elements that were finalized in the 2016 final rule. The comment period on that NPRM closed on June 18, 2019.

Feasibility to Link to Other Data

Personally identifiable information:

There is no PII sufficient to be used for linking purposes. Encrypted identifiers are included, which allow for linking over time and with the National Child Abuse and Neglect Database (NCANDS) and National Youth in Transition Database (NYTD) records. Exact dates of birth and other date information that could be used for identification are collected, though only perturbed versions are made available in datasets distributed by NDACAN.

Geographic and institutional identifiers:

State and county identifiers are available in the datasets available through NDACAN. The county identifier represents the county that has responsibility for the care of the child, rather than the child's residence. County identifiers are suppressed if there are fewer than 1,000 records reported. NDACAN is examining alternatives for making county identifiers available through a restricted access virtual data center. In-house datasets contain all county identifiers, but access is limited to federal staff.

Data it has been linked to in the past:

There is a unique person-level identifier that allows researchers to link records across years within AFCARS, and allows for linking to individual child records reported to NCANDS. They may also be linked to person records in NYTD, a database containing information on 17-year-old youth transitioning out of foster care. The person records are unique only within a state, however, so children cannot be followed across states. All three (AFCARS, NCANDS, and NYTD) are available for analysis through NDACAN.

Example research products using linked data:

AFCARS foster care linked with NYTD:

Shpiegel, S., Cascardi, M., & Dineen, M. (2016). A social ecology analysis of childbirth among females emancipating from foster care. *Journal of Adolescent Health*. https://doi.org/10.1016/j.jadohealth.2016.09.013

Shpiegel, S., & Ocasio, K. (2015). Functioning patterns among older adolescents in foster care: Results from a cluster analysis. *Children and Youth Services Review*, 58, 227–235. https://doi.org/10.1016/j.childyouth.2015.09.024

Okpych, N. J. (2015). Receipt of independent living services among older youth in foster care: An analysis of national data from the U.S. *Children and Youth Services Review*, 51, 74–86. https://doi.org/10.1016/j.childyouth.2015.01.021

AFCARS linked with NCANDS:

Billings, P., & Moore, T. (2004). Child Maltreatment in Foster Care (Outcomes research report No. 8). Oklahoma Department of Human Services. Retrieved from

http://cornell.worldcat.org/title/child-maltreatment-in-foster-care/oclc/608436109

Boles, S. M., Young, N. K., Dennis, K., & DeCerchio, K. (2012). The regional partnership grant (RPG) program: Enhancing collaboration, promising results. *Journal of Public Child Welfare*, 6(4), 482–496. https://doi.org/10.1080/15548732.2012.705239

Conn, A.-M., Szilagyi, M. A., Franke, T. M., Albertin, C. S., Blumkin, A. K., & Szilagyi, P. (2013). Trends in child protection and out-of-home care. *Pediatrics*, 132(4), 712–719. https://doi.org/10.1542/peds.2013-0969

Hill, R. B. (2007). An Analysis of Racial/Ethnic Disproportionality and Disparity at the National, State, and County Levels. Casey-CSSP Alliance for Racial Equity in Child Welfare. Retrieved from https://www.aecf.org/resources/an-analysis-of-racial-ethnic-disproportionality-and-disparity-at-the-nation/

Kaftan, J. (2016). National identity and governmental authority: The intersection of national identity, immigration policy, and the child welfare system in the United States. In B. Wejnert & P. Parigi (Eds.), *Research in Political Sociology* (Vol. 24, pp. 67–92). Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/doi/10.1108/S0895-993520160000024004

Lancaster, L., & Fong, R. (2014). Disproportionality and Disparities for Latinos and Native Americans: Predictors of Race/Ethnicity in Decision Points in Public Child Welfare. In The Society for Social Work and Research 2014 Annual Conference. San Antonio, TX. Retrieved from https://sswr.confex.com/sswr/2014/webprogram/Paper21223.html

McDonald, T., Billings, P., & Moore, T. (2002). Comparison of Oklahoma's 2000 AFSA Outcomes by Administrative Area (Outcomes research report No. 2). Oklahoma City, OK: Oklahoma Department of Human Services. Retrieved from http://digitalprairie.ok.gov/cdm/ref/collection/stgovpub/id/14504

McDonald, T., & Testa, M. F. (2010). Outcomes Monitoring in Child Welfare. In M. F. Testa & J. Poertner (Eds.) (pp. 101–135). Oxford Univ Pr. Retrieved from https://books.google.com/books?id=p2QSDAAAQBAJ&printsec=frontcover&dq=M.+F.+Testa+%26+J.+Poertner+(Eds.)+2010+OUTCOMES+MONITORING+CHILD+WELFARE&hl=en&sa=X&ved=0ahUKEwjykaji4-jgAhWluVkKHerIDqoQ6AEIKDAA#v=onepage&q=M.%20F.%20Testa%20%26%20J.%20Poertner%20(Eds.)%202010%20OUTCOMES%20MONITORING%20CHILD%20WELFARE&f=false

McKlindon, A., Vandivere, S., & Murphey, D. (2011). Foster care data snapshot. *Child Trends*, 19, 1–7. Retrieved from https://www.childtrends.org/wp-content/uploads/2013/03/Child_Trends_2011_05_31_DS_FosterCare.pdf

National Data Analysis System. (2004). Children of Color in the Child Welfare System.

Washington, DC: Child Welfare League of America. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwii5MLsrtrOAhXClR4KHUYBBzcQFggcMAA&url=http%3A%2F%2Fwww.ibrarian.net%2Fnavon%2Fpaper%2FChildren_of_Color_in_the_Child_Welfare_System.pdf%3Fpaperid%3D1223873&usg=AFQjCNH2VIBcojQjHcbTyk-4U-JMl141qw&sig2=UxPLSqBicShgpUgV33mjxg&bvm=bv.129759880,d.dmo

Schuerman, J. R., & Needell, B. (2009). The Child and Family Services Review Composite Scores: Accountability off the Track. Chicago, IL: Chaplin Hall at the University of Chicago.

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2001a). Child and Family Services Review. Final Assessment: North Carolina. Washington, DC: North Carolina State Department of Health and Human Services.

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2001b). Child and Family Services Reviews. Final Assessment: Minnesota. St. Paul, MN: Minnesota State Department of Health and Human Services.

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2006). Targeted Grants to Increase the Well-Being of, and to Improve the Permanency Outcomes for, Children Affected by Methamphetamine or Other Substance Abuse: First Annual Report to Congress. Retrieved from http://www.acf.hhs.gov/programs/cb/resource/targeted-grants-to-increase-the-well-being

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2008). Child Welfare Outcomes 2002-2005: Report to Congress. Washington, DC: U.S. Department of Health and Human Services, Administration for Children, Youth and Families. Retrieved from http://www.acf.hhs.gov/programs/cb/pubs/cwo05/index.htm

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2010a). Child Welfare Outcomes 2003-2006: Report to Congress. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. Retrieved from http://www.acf.hhs.gov/programs/cb/pubs/cwo03-06/cwo03-06.pdf

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2010b). Child Welfare Outcomes 2004-2007: Report to Congress. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. Retrieved from www.acf.hhs.gov/programs/cb/pubs/cwo04-

07/cwo04-07.pdf

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2012). Targeted grants to increase the well-being of, and to improve the permanency outcomes for, and children affected by methamphetamine or other substance abuse: third annual report to congress. Retrieved from http://www.cffutures.org/files/RPG%20Program_Third%20Report%20to%20Congress.pdf

U.S. Department of Health and Human Services; Administration for Children and Families; Children's Bureau. (2013). Targeted grants to increase the well-being of, and to improve the permanency outcomes for, and children affected by methamphetamine or other substance abuse: fourth annual report to congress: Representing the initial grant period of September 30, 2007, to September 30, 2012.

(All citations from NDACAN's CanDL.)

Potential Research Areas to Inform Policy

Linking AFCARS and NCANDS to Medicaid data, data from courts, or housing data could yield greater insight into the populations being served, and what other efforts may be made to better address their needs and provide services to prevent foster care placement and support children in their homes. Due to the lack of PII in AFCARS, however, such research based on linked data would need to rely on data from state agencies.

Data Quality

Validity and reliability:

The validity and reliability of data reported to AFCARS is assessed formally through the AFCARS Assessment Reviews. The complete guide to this assessment can be found at https://www.acf.hhs.gov/sites/default/files/cb/afcars_assessment_review_guide.pdf

Nonresponse and coverage:

Because AFCARS is required by law, and items are collected uniformly by state child welfare agencies and not based on individual response, item non-response is not as much of an issue. Some data elements have been more prone to underreporting than others; for example, some caseworkers may be likely to mark only "neglect" as a circumstance associated with removal from home and not mark all circumstances that apply, but this issue has not been observed as much in recent years. Coverage should be complete as agencies are required to enter data on all children for whom they have care and placement responsibility.

Documentation:

See AFCARS Codebook:
https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/AFCARSFosterCareCodebook.pdf

User Guide:
http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/AFCARSGuide2000toPresent.pdf

Overall quality:

The overall data quality of AFCARS is considered very high by subject matter experts, partly because it serves as a major source for federal assessment of state performance, and it is required reporting that has been in existence since 1995, which has allowed for remarkable improvement in data quality. There are, however, notable variations in data quality by state and by topical area. Changes to reporting in AFCARS may begin in 2020, which may introduce new issues around data quality and reporting. Currently, data quality while very good as a whole varies from state to state and by element. Some elements are more reliable because of frequent use, such as discharge reasons, dates of discharge, latest removal for foster care episodes, and dates of birth. Other elements are not as quality controlled or used at the federal level for assessing outcomes, such as the completeness of circumstances associated with removal and disability information. Certain states and territories, such as Puerto Rico, have long struggled to meet data quality standards, so their data must be interpreted with caution.

Program Scope and Budget

Size of population served annually:

A total of 443,000 children in foster care on September 30, 2017; 691,000 total served in FY 2017. There were 59,400 children adopted with child welfare agency involvement in FY 2017. Source:
https://www.acf.hhs.gov/sites/default/files/cb/trends_fostercare_adoption_08thru17.pdf

Federal funding for programs encompassed by the data:

State FY 2016: $2.667 billion from Title IV-E Foster Care Program. [Source: Child Trends, (2018) Child Welfare Financing SFY 2016]. https://www.childtrends.org/wp-content/uploads/2018/12/CWFSReportSFY2016_ChildTrends_December2018.pdf]

Child Care and Development Fund (CCDF), Office of Child Care Information System (OCCIS) – Case Level Administrative Data (ACF-801 data)

Website

https://www.acf.hhs.gov/occ/resource/ccdf-statistics and
http://www.researchconnections.org/childcare/series/00215

Funding Agency and Data Ownership:

Office of Child Care (OCC), ACF/HHS

Data Gathering Agency/Contractor

OCC, under contract with General Dynamics Information Technology

Overview

When Congress created the CCDF it also created the requirement that case-level data on families receiving CCDF services be collected on a monthly or quarterly basis. States and territories were charged with submitting specific information. Although “ACF-801” is only a number that the OMB assigned to the data collection form that ACF designed to meet the statutory reporting requirements, the term is used today to refer to case-level CCDF data.

Refer to https://www.acf.hhs.gov/occ/resource/overview-of-case-level-reporting-for-states-and-territories-acf-801 for more details.

All lead agencies in the states, the District of Columbia, and territories (including Puerto Rico, American Samoa, Guam, Northern Mariana Islands, and the U.S. Virgin Islands) are responsible for collecting and reporting ACF-801 data. States/territories can submit either full population or sample data.

Guidance to states/territories regarding sampling is available at https://www.acf.hhs.gov/sites/default/files/occ/tb5r_v2.pdf.

Periodicity

The cases are monthly data (e.g., there are cases for January, separate cases for February, etc.). The data are released to the public on an annual basis.

Data Content and Structure

In general, states/territories report a summary record for their data submission, including the number of families served.

For families, data elements include single-parent status of the head of household, monthly co-payment amount, reasons for care (e.g., employment, training/education, protective services, etc.), income used to determine eligibility, family size on which eligibility is based, homeless status (new element), family ZIP code (new element), and primary language spoken at home (new element).

For children, data elements include ethnicity, race, gender, date of birth, and child disability (new element).

For settings (environment in which the child received care), data elements include the type of child care setting, the total amount paid to the provider, and the total number of hours of care received by the child.

For providers, data elements include provider QRIS participation, provider QRIS rating, provider accreditation status, provider ZIP code, provider subject to state pre-K standards, and provider subject to Head Start/Early Head Start standards.

In addition, the OCC collects the pooling factor (through the annual aggregate ACF-800 form), which indicates if the state/territory reported only children served with CCDF funds (100 percent pooling factor), or children served with CCDF and other funding sources (less than 100 percent pooling factor). The pooling factor indicates the percentage of child care funds that is provided through the CCDF.

Major Publications

Office of Child Care Reports to Congress are available at https://www.acf.hhs.gov/occ/resource/reports-to-congress

“Estimates of Child Care Eligibility and Receipt for Fiscal Year 2015” is available at https://aspe.hhs.gov/system/files/pdf/260361/CY2015ChildCareSubsidyEligibility.pdf

“Child Care Subsidy Duration and Caseload Dynamics: A Multi-state Examination” is available at https://aspe.hhs.gov/system/files/pdf/77151/rpt_ChildCareSubsidy.pdf

“Researching the CCDF Program by Linking Administrative Data with Data from the CCDC Policies Database: A How-To Guide” is available at https://www.acf.hhs.gov/opre/resource/researching-ccdf-linking-admin-data-ccdf-policies-database-how-to-guide

“Employment Outcomes for Low-Income Families Receiving Child Care Subsidies in Illinois, Maryland, and Texas” is available at http://www.researchconnections.org/childcare/resources/16511/pdf (this study uses administrative data directly from the states, rather than the national data compiled by OCC).

Available Datasets

Public-use aggregate:

https://www.acf.hhs.gov/occ/resource/ccdf-statistics and http://www.researchconnections.org/childcare/series/00215 (after selecting a year, scroll down to DS6, DS7, DS8, and DS9 – these links allow users to create simple crosstabs and frequency tables).

Public-use micro:

http://www.researchconnections.org/childcare/series/00215

Restricted-use micro:

There is no version of the ACF-801 child care administrative data at this level of restricted use.

In-house micro:

In-house data contain the full population of cases for states that choose to submit population (rather than sample) data. Also included are identifiers that allow cases to be linked across time. Prior to the Child Care and Development Block Grant (CCDBG) Act of 2014, PII were included in these data.

HHS has collaborated with Census’ Center for Administrative Records Research and Applications (CARRA) to link to Census data and other administrative data held by CARRA. Federal partners may access the CARRA data onsite at Census. Researchers may access the CARRA data only after being selected through a competitive proposal process with Chapin Hall.

Statutory and Regulatory Restrictions

Privacy Act, System of Records Notice. In addition, the reauthorization of the CCDBG states that data submitted to the OCC shall not contain PII.

Feasibility to Link to Other Data

Personally identifiable information:

ACF-801 data collected prior to October 2015 contain SSNs (optional field) and state/territory case unique identifiers. After the re-authorization of the CCDBG, states and territories no longer report SSNs, but they do provide state/territory case unique identifiers that could potentially allow for linking records over time.

Geographic and institutional identifiers:

System-generated unique identifiers are available in the public-use data.

Data it has been linked to in the past:

The data have been linked to the American Community Survey (ACS). Data from three states (Illinois, Maryland, and Texas) have also been linked to the ACS, Unemployment Insurance (UI) wage report data, Temporary Assistance for Needy Families (TANF) program data, and Longitudinal Employer-Household Dynamics (LEHD) data.

Example research products using linked data:

"Employment Outcomes for Low-Income Families Receiving Child Care Subsidies in Illinois, Maryland, and Texas" is available at http://www.researchconnections.org/childcare/resources/16511/pdf (this study uses administrative data directly from the states, rather than the national data compiled by the OCC).

"Is Subsidized Childcare Associated with Lower Risk of Grade Retention for Low-Income Children? Evidence from Child Care and Development Fund Administrative Records Linked to the American Community Survey" (conference proceedings) is available at https://www.census.gov/fedcasic/fc2017/abstracts/1_9_2.html.

Potential Research Areas to Inform Policy

Data have been used to study caseload dynamics and subsidy costs, which can be used to understand differences across states and relating those differences to policies and economic context. ASPE makes available data on CCDF subsidy spells, which researchers can use for these and other purposes: https://aspe.hhs.gov/pdf-report/child-care-subsidy-duration-and-caseload-dynamics-multi-state-examination-2004-2014.

Linking this child care data with employment data (e.g., ACS, CPS, and UI data) could answer questions about the relationship between subsidy receipt and employment outcomes. Longitudinal (panel) employment data would be best.

Linking subsidy data with provider wage data could allow the exploration of the effects of subsidy amount on provider wages. Linking child care data with a dataset that contains quality indicators (e.g., state-held data) could answer interesting questions about the relationship between aspects of subsidy receipt and provider quality.

Linking with CPS data could answer questions about child care expenditures of families receiving and not receiving subsidies.

Data Quality

Validity and reliability:

ASPE is currently undertaking a project to compare variables from the child care data with the ACS data (e.g., income is available in both datasets, so after cases are linked, it's possible to compare 801-reported income and ACS-reported income).

Payment (Cost): In FY 2011, OCC changed the reporting requirement for this data element. Whereas this data element was previously defined as the full amount paid to the provider each month (CCDF subsidy plus copay), beginning in FY 2011, this data element was limited to only the subsidy amount paid to the provider.

Many states and families have apparent inconsistencies between the Reason for Receiving Subsidized Child Care, Total Income for Determining Eligibility, and Sources of Income. These inconsistencies may be real inconsistencies, or they may be the result of state definitions and policies. For example, a family reports employment as a reason for care, but has zero income and indicates that employment is not a source of income. It might be that there is a true inconsistency between the data elements; or it could be that the state definition of employment includes looking for work (job search), in which case the reason for care and income variables are not truly inconsistent.

There also appear to be inconsistencies between Single Parent, Reason for Receiving Subsidized Child Care, and Family Size. Again, these inconsistencies may reflect true error, or may reflect differences in policies regarding how states define, code, and extract protective services and foster care cases (from setting records codebook).

Nonresponse and coverage:

Coverage is good and allows for analysis at the state level (a minimum of 2,400 cases per state/territory). Item non-response is high for the Race and Ethnicity variables. This dataset does not contain any tribal child care data. That information is contained within another aggregate level dataset.

Documentation:

Detailed documentation can be found at http://www.researchconnections.org/childcare/series/00215.

Overall quality:

Data quality is generally considered high by subject matter experts. Income from one or two territories in one or two years was misreported as annual income, when it should have been monthly income.

Program Scope and Budget

Size of population served annually:

On average, 1.3 million children were served per month in FY 2017 (preliminary data).

Federal funding for programs encompassed by data:

Federal expenditures were $6.5 billion in FY 2016.

Federal Case Registry (FCR)

Website

https://www.acf.hhs.gov/css/resource/federal-case-registry-information-for-families

Funding Agency and Data Ownership

Office of Child Support Enforcement (OCSE), ACF/HHS

Data Gathering Agency/Contractor

OCSE, with support of federal contractors. Data are maintained at the Social Security Administration's National Computer Center.

Overview

The FCR is a national database that includes all child support cases handled by state child support agencies (referred to as IV-D cases), and all support orders established or modified on or after October 1, 1998 (referred to as non IV-D orders). It contains state child support (IV-D) and non-IV-D case data and serves as a pointer system to help locate persons across state lines. Person data in the FCR are matched daily against new employment data received in the National Directory of New Hires (NDNH) and sent to states to facilitate case processing and increase collections, especially through automated income withholding. Matches are sent to states to inform them if a IV-D case participant in their state appears as a participant in a IV-D or non-IV-D case in another state.

The FCR also serves as the conduit for matching against the following sources: Department of Defense, Department of Veterans Affairs, Federal Bureau of Investigation, Internal Revenue Service, National Security Agency, and Social Security Administration (SSA). Matches made through the Multistate Financial Institution Data Match, Thrift Savings Plan match, and the Insurance Match are returned to states through the FCR, as well as medical coverage matches made with the Defense Enrollment Eligibility Reporting System maintained by the Defense Manpower Data Center.

Periodicity

A state has the option of selecting its own schedule for transmission of data to the FCR. The FCR is a batch system that executes on a daily basis, Monday through Friday.

Data Content and Structure

Information in the FCR is composed of extracts from each state’s State Case Registry, including data on persons attached to a case such as the custodial parent, the noncustodial parent, the putative father, and the child. A family violence indicator can also be placed on the data, restricting the sharing of any information related to the case.

Major Publications

FY 2016 Annual Report to Congress. Office of Child Support Enforcement. Page 8. https://www.acf.hhs.gov/css/resource/fy-2016-annual-report-to-congress

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

The FCR is a restricted use dataset that includes PII. Each state’s statewide automated child support enforcement system exchanges information with the FCR. Detailed information on who can access the data, and under what circumstances, can be found in Charts 1-2 in the above link. (See https://www.acf.hhs.gov/css/resource/federal-case-registry-interface-guidance-document.)

In-house micro:

Select federal agencies may access the FCR dataset for specific locating purposes outlined in statute (Section 453 and 463 of the Social Security Act). Additionally, “the Secretary may provide access to data in each component of the FPLS... for research purposes found by the Secretary to be likely to contribute to achieving the purposes of part A or this part, without personal identifiers.” See https://www.ssa.gov/OP_Home/ssact/title04/0453.htm.

Statutory and Regulatory Restrictions

Section 453 (l)(m) and (b)(2) of the Social Security Act contain specific references to security and privacy requirements at the federal level. In addition, States are required to have policies and procedures in place to monitor access, transmit data to the federal level, and maintain safeguards and system controls.

Feasibility to Link to Other Data

Personally identifiable information:

The FCR database consists of tables that include PII. Some of the basic tables that include PII are: FCR Case Data, including state/territory code, case ID, case type, and court order indicator; Person Data, including SSN, name, sex, DOB, and family violence indicator; and Other Person Identifiers, including name (if any), SSN, sex, and the DOB of other persons associated with the person’s child support case.

Geographic and institutional identifiers:

Geographic identifiers such as state/territory code are included in the FCR.

Data it has been linked to in the past:

The FCR database is proactively matched with the Federal Parent Locator System (FPLS), not for research purposes but to help locate persons across state lines. Person data in the FCR are matched daily against new employment data received in the NDNH and sent to states to facilitate case processing and increase collections, especially through automated income withholding. Matches are sent to states to inform them if a IV-D case participant in their state appears as a participant in a IV-D or non-IV-D case in another state.

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

Linking the FCR data to Census data would give policymakers a sense of the composition of the IV-D caseload. Currently, characteristics of the IV-D caseload are based on formulas derived from survey data. Additionally, linking the FCR and NDNH data to the Census’ LEHD project data would enable policymakers to get a sense of noncustodial parents’ employment and their employers. See https://lehd.ces.census.gov/data/.

Data Quality

The FCR dataset is generally regarded as high quality. SSN/name combinations are submitted to the SSA for verification when sufficient data exists to support the verification process.

Validity and reliability:

See https://www.acf.hhs.gov/sites/default/files/ocse/fcr_igd_app_e_data_dictionary.pdf.

Nonresponse and coverage:

There are no known gaps in the data collection. Nonresponse for a variable such as current address may impede locate functions.

Documentation:

OCSE publishes an Interface Guidance Document. It also supports states with ongoing technical assistance and has liaisons on staff to respond to questions.

Overall quality:

The data quality of the FCR dataset is generally considered high by subject matter experts.

Program Scope and Budget

Size of population served annually:

The number of unique (verified) people in the FCR in FY 2016 was 45.4 million. Source: https://www.acf.hhs.gov/sites/default/files/programs/css/fy_2016_annual_report.pdf, Table 97.

Federal funding for programs encompassed by data:

Total expenditures for the child support program in FY 2016 were $5.7 billion ($3.4 billion in federal funds and $2.3 billion in state funds).

Low Income Home Energy Assistance Program (LIHEAP) data, Appended to the Residential Energy Consumption Survey (RECS)

Website

https://www.acf.hhs.gov/ocs/resource/2015-recs-liheap-household-administrative-data-matching

Funding Agency and Data Ownership

Division of Energy Assistance; Office of Community Services. ACF/HHS

Data Gathering Agency/Contractor

Applied Public Policy Research Institute for Study and Evaluation (APPRISE)

Overview

RECS is a periodic study conducted by the U.S. Energy Information Administration (EIA) that provides detailed information about energy usage in U.S. homes. The household survey collects data on energy-related characteristics and usage patterns of a national representative sample of housing units. The Energy Supplier Survey (ESS) collects data on how much electricity, natural gas, propane/LPG, fuel oil, and kerosene were consumed in the sampled housing units during the reference year. It also collects data on actual dollar amounts spent on these energy sources.

More details on the RECS are available at https://www.eia.gov/consumption/residential/reports/2015/methodology/.

Through a data matching effort, OCS collects LIHEAP administrative data records from states and links them to records in the RECS survey.

Periodicity

Approximately every four to six years.

Data Content and Structure

LIHEAP

- Name
- Household address
- Amount of LIHEAP benefits

- Household income
- Household name
- ID from IT system
- Household size (number of household members)
- Type of LIHEAP assistance received
(heating, cooling, crisis, other payment assistance)
- Breakdown of amount of LIHEAP benefits by type of assistance received
(heating, cooling, crisis, other bill payment assistance)
- Date(s) of LIHEAP assistance by type of LIHEAP assistance
(heating, cooling, crisis, other bill payment assistance)
- Presence of young child member (age 5 or younger) in household
- Presence of elderly member (60+) in household
- Presence of disabled member in household
- Tenancy of home (own or rent)
- Type(s) of fuel used in home
- Indicator whether heating expenditures are included in rent

RECS
All household survey questionnaires covered the same topical areas, with questions about the type and number of energy-consuming devices, usage patterns, structural characteristics of the home, household demographics, and energy supplier information.

The questionnaire topics were as follows:
- Housing unit characteristics
- Appliances
- Electronics
- Space heating
- Air conditioning
- Water heating
- Lighting
- Energy programs
- Energy bills
- Energy suppliers
- Household characteristics
- Energy assistance
- Housing unit measurement (for in-person interview only)
- Scanning of sample energy bills (for in-person interview only)

Major Publications

LIHEAP Home Energy Notebook: https://www.acf.hhs.gov/ocs/resource/liheap-home-energy-notebooks

LIHEAP report to Congress: https://www.acf.hhs.gov/ocs/resource/liheap-reports-to-congress

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

None. The dataset, though free of conventional PII, has sensitive information the EIA does not release it.

In-house micro:

The in-house dataset is available only to the contractor. If it were to be made available to those outside, it would be made available only to those with authority to access it under the Confidential Information Protection and Statistical Efficiency Act (CIPSEA).

Statutory and Regulatory Restrictions

The LIHEAP block grant (42 U.S.C. 8621) was established under Title XXVI of the Omnibus Budget Reconciliation Act of 1981, Public Law 97-35. The law requires the collection of data on eligible households.

Details can be found in section 2610 at https://www.acf.hhs.gov/ocs/resource/liheap-statute-and-regulations.

Feasibility to Link to Other Data

Personally identifiable information:

PII are collected to link to RECS survey records and then destroyed. PII collected includes name, address, and ID from IT system.

Geographic and institutional identifiers:

State and state groups (for less populated states). With the reduction in sample size for the 2015 survey, it is not clear whether any state information will be offered.

Data it has been linked to in the past:

RECS.

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

A link to Supplemental Nutrition Assistance Program (SNAP) data would be useful. Some SNAP allotments are tied to energy assistance receipt.

Data Quality

Validity and reliability:

Some internal consistency checks on the data are performed. Consistency is considered pretty good. Two states still cannot offer unduplicated counts for “any” receipt.

Nonresponse and coverage:

The RECS is a nationally representative survey of housing units, and LIHEAP data are appended to those records.

Documentation:

None that is publicly available.

Overall quality:

For overall counts, the RECS data are considered of high quality by subject matter experts. The quality of income data in RECS is considered variable. Disability is not consistently defined across states, and, therefore, RECS data on this aspect should be used with caution.

Program Scope and Budget

Size of population served annually:

Over 5 million served in 2013. Source: https://liheapch.acf.hhs.gov/docs/LIHEAPprimer.pdf.

Federal funding for programs encompassed by data:

Over $4.5 billion in 2013. Source: https://liheapch.acf.hhs.gov/docs/LIHEAPprimer.pdf.

National Child Abuse and Neglect Data System (NCANDS)

Website

https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/ncands

Funding Agency and Data Ownership

Children's Bureau, ACYF/ACF/HHS

Though data are owned by the Children's Bureau, it is set up as a voluntary data collection effort and states are given a great deal of control over the data.

Data Gathering Agency/Contractor

Walter R McDonald & Associates, Inc. (WRMA)

Overview

NCANDS is a federally sponsored effort that annually collects and analyzes data on child abuse and neglect known to child protective services (CPS) agencies in the United States. The mandate for NCANDS is based on the 1988 amendments to the Child Abuse Prevention and Treatment Act (CAPTA), which directed the HHS Secretary to create a national data collection and analysis program for state-level child abuse and neglect information. Subsequent amendments to CAPTA have led to new data collection requirements, many of which are incorporated into NCANDS.

State participation in NCANDS is voluntary. Thus, an essential component of NCANDS is a collaborative working partnership between the federal government and the states and territories. Every state submits two files annually –a Child File containing case-level records for each report of alleged child abuse and neglect that received a CPS response, and an Agency File containing aggregate data. The data submissions are reconfigured into relational files for analysis.

A validation process occurs for all data submission files using code checking and inter- and cross-record rule checking to ensure compliance with the federal government's and NCANDS' data standards. A clean dataset is loaded into a multidimensional database and analyses are conducted. The analyses are publicly released in the annual Child Maltreatment report series. Child Maltreatment 2015 is the 26th report released by NCANDS on behalf of ACYF. NCANDS data also are a critical source of information for many publications, reports, child welfare personnel, researchers, and others. NCANDS data are used to measure the performance of several federal programs and are an integral part of the Child and Family Services Reviews (CFSR) and the Child Welfare Outcomes: Report to Congress. See https://www.acf.hhs.gov/cb/resource/about-ncands.

Periodicity

Annual.

Data Content and Structure

Child File

- Report Data (fields 1–11): The two identifying fields (submission year and state ID) and general information about the report.
- Child Data (fields 12–25): General information about the specific child in the record. All fields in this section are attributes related to the child ID.
- Maltreatment Data (fields 26–34): Information about maltreatment types and maltreatment disposition levels.
- Child Risk Factors (fields 35–43): Data about the child's characteristics or environment that may place the child at risk for maltreatment. This includes diagnosed disabilities, alcohol and drug abuse, and behaviors or problems.
- Caregiver Risk Factors (fields 44–55): Data about the child's caregiver characteristics or environment that may place the child at risk for maltreatment. This includes domestic violence, substance abuse, and financial problems.
- Services Provided (fields 56–85): Information about services that are provided for the child or family.
- Staff Data (fields 86–87): Identification information about the CPS worker and the CPS worker's supervisor who were associated with the child on the date of the report disposition.
- Perpetrator Data (fields 88–144): Information about perpetrators of maltreatment.
- Additional Fields (fields 145–150): Any new fields that were added to the Child File subsequent to its creation in 2001.

Agency File

- Preventive Services (fields 1.1.A–1.1.E): Data on recipients of preventive services by various funding sources.
- Additional Information on Referrals and Reports (fields 2.1–2.4): Information on the number of referrals and children with child abuse and neglect allegations, but the referrals were screened out prior to being referred for investigation or assessment, response time with respect to the initial investigation or assessment, workforce, and caseload.
- Additional Information on Child Victims Reported in the Child File (fields 3.1–3.5): Data about family preservation and reunification services, and out-of-court contacts for victims reported in the Child File.
- Information on Child Fatalities Not Reported in the Child File (fields 4.1–4.4): Data for maltreatment deaths that are not already reported in the Child File; this

includes maltreatment deaths that occurred in foster care, after family preservation services, and after family reunification services.

- Individuals with Disabilities Education Act (IDEA) (fields 5.1–5.2): Data for children eligible for referral and children referred to agencies providing early intervention services under IDEA Part C.

Major Publications

Child Maltreatment: https://www.acf.hhs.gov/cb/research-data-technology/statistics-research/child-maltreatment

Available Datasets

Public-use aggregate:

Findings from the National Child Abuse and Neglect Data System (NCANDS) – Data Tables (updated with FY 2015), which contains state-level counts, percentages, and rates for selected measures, is available at https://www.healthdata.gov/dataset/national-child-abuse-and- neglect-data-system-ncands-child-file.

Public-use micro:

None.

Restricted-use micro:

NDACAN makes restricted data available to eligible members of the research community for the purpose of statistical analysis and reporting. Restricted datasets are licensed for a three-year period which is renewable. Researchers who would like to use restricted data must fulfill eligibility criteria, and submit and enter into a legally binding data license that outlines the requirements for appropriate use of the restricted data. For additional information, visit http://www.ndacan.acf.hhs.gov/datasets/request-restricted-data.cfm. Privacy is protected using a number of techniques, including rounding of dates; suppression or exclusion of certain geographic, date, and ID variables; and double encryption of unique identifiers.

In-house micro:

In-house datasets include encrypted state-specific person identifiers so that individuals can be followed over time, and in other federal child welfare datasets (AFCARS and NYTD). No data are suppressed or altered in the in-house version.

Statutory and Regulatory Restrictions

States provide unique encrypted person ID numbers for each child in their data submission, but no PII that would allow for linking with other datasets except for

AFCARS and NYTD (see elsewhere in this volume). There does not appear to be any statutory prohibition, however, on the reporting of PII that could be used for linking to other non-child welfare datasets.

NCANDS was established in response to the 1988 amendment (P.L. 100-294) to CAPTA (42 U.S.C. 5101 et seq.), which called for the creation of a coordinated national data collection and analysis program, both universal and case-specific in scope, to examine standardized data on false, unfounded, or unsubstantiated reports. The 1996 CAPTA amendment (42 U.S.C. 5106a(d)) required all states that receive basic state grant funds to provide specific data elements, to the extent practicable, to the federal government. These data items were incorporated into NCANDS. Since that time, other CAPTA revisions that pertain to child welfare have been included in NCANDS. The most recent amendments to CAPTA during 2015 and 2016 (P.L. 114-22 and P.L. 114-198) retained and expanded upon those provisions.

Feasibility to Link to Other Data

Personally identifiable information:

There is no PII sufficient to be used for linking purposes. Encrypted identifiers are included, which allow for linking over time and with AFCARS and NYTD records.

No identifying data, such as name, address, or SSN, are collected. The only PII collected are dates of birth, military status indicator, and military family member indicator. NCANDS does not collect information directly from individuals; all information is received from state child welfare agencies. NCANDS underwent the Authorization to Operate (ATO) security review process and received approval for ATO on April 18, 2017. To ensure the confidentiality of the Child File data, each state encrypts its identifiers. No actual case or individual identifiers are submitted. Each state ensures that its data meet a standard of encryption. Before the data are released for restricted public use through NDACAN at Cornell University, the Child File identifiers are double-encrypted, adding another layer of confidentiality.

Geographic and institutional identifiers:

State and county identifiers are available in the datasets available through NDACAN. The county represents the one that has responsibility for the care of the child, rather than the child’s residence. County identifiers are suppressed if there are fewer than 1,000 records reported. NDACAN is examining alternatives for making county identifiers available through a restricted access virtual data center. In-house datasets contain all county identifiers, but access is limited to federal staff.

Data it has been linked to in the past:

There is a unique person-level identifier that allows researchers to link records across years within NCANDS, and allows for linking to foster care and adoption records in

AFCARS. They may also be used to link to person records in the National Youth Transition Database (NYTD), a database containing information on 17-year-old youth transitioning out of foster care. The person records are unique only within state, however, so children cannot be followed across state. Finally, NCANDS-like records collected directly from the states have been linked to the National Survey of Child and Adolescent Well-Being (NSCAW), and is available for analysis through NDACAN. At the state and local levels, NCANDS-like records have been incorporated into ongoing integrated data systems (IDS) in a number of places.

Example research products using linked data:

Wilderman C, Emanuel, M, Levinthal, J, Putnam-Hornstein, E, Waldfogel, J, and Hedwig, L. (2014). *The Prevalence of Confirmed Maltreatment Among American Children, 2004-2011*. JAMA Pediatr. 168(8): 706-713.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5087599/

Potential Research Areas to Inform Policy

- Linked records across NCANDS, AFCARS, and NYTD data can allow tracking child outcomes through the course of their involvement with child welfare systems, identifying patterns in length of stay and relationships among various child and case characteristics.
- Connecting state data that feed into NCANDS to other data systems, such as national child death data, would allow for the development of predictive models that could lead to early identification and effective interventions for high-risk families. Currently NCANDS does not collect PII from states.
- Connecting state data that feed into NCANDS to employment data would allow researchers to better model the effects of maltreatment on a successful transition to adulthood, and possibly identify child welfare program supports that lead to better long-term outcomes for maltreated children.

Data Quality

Validity and reliability:

CAPTA legislation recognizes individual state authority by providing a minimum federal definition of child abuse and neglect. Each state defines child abuse and neglect in its own statutes and policies and the child welfare agencies determine the appropriate response for the alleged maltreatment based on those statutes and policies. NCANDS was created with the objective to collect nationally standardized aggregate and case-level data and to make these data useful for policy decision-makers, child welfare researchers, and practitioners. NCANDS developed a general mapping procedure whereby all states could systematically define the rules for mapping the state data fields and codes and define the rules for extracting the data from the state system into the standard NCANDS data format. NCANDS Technical Team members provide one-on-one technical assistance to states to assist with data mapping, construction, extraction, and reporting.

During the development of NCANDS, submission file restructuring, and whenever new data elements are proposed, state CPS agency representatives are involved with determining how the data will be collected, variable definitions, and validation rules. A State Advisory Group of 25 states helped with the initial design and development of NCANDS, and State Working Groups of nine states continue to help with adding new data elements.

Nonresponse and coverage:

As mandated by the CAPTA legislation, NCANDS collects data about child abuse and neglect incidents that are known to CPS agencies during each federal fiscal year. Because NCANDS contains all screened-in referrals to CPS agencies that received a disposition, including those that received an alternative response, these data represent the universe of known CPS child maltreatment cases.

NCANDS has a standard data format for state submission of child abuse and neglect data. In contrast to this standard, each state has a unique system for gathering and reporting data. The states also vary greatly in the policies, legislation, requirements, and processing methods for the collection and storage of the data at the state level. NCANDS developed a general mapping procedure whereby all states could systematically define the rules for mapping the state data fields and codes and define the rules for extracting the data from the state system into the standard NCANDS data format. A successful federal-state partnership is the core component of NCANDS. Each state designates one person to be the NCANDS state contact. The NCANDS state contacts from all 52 states work with the Children’s Bureau and the NCANDS Technical Team to uphold the high-quality standards associated with NCANDS data. Webinars, technical bulletins, virtual meetings, e-mail, listserv discussions, and phone conferences are used regularly to facilitate information sharing and provision of technical assistance.

Principal variables have very high data quality. Data gaps tend to be in areas that are difficult to accurately assess or measure and may go undetected in the child or caregiver. Risk factors are one such area, especially as some risk factors must be clinically diagnosed, which may not occur during the investigation or alternative response. If the child maltreatment case is closed prior to the diagnosis, the CPS agency may not be notified of the diagnosis and the information will not be reported to NCANDS. Another area is services provision. NCANDS collects data for 26 types of services including adoption, employment, mental health, and substance abuse. States have their own typologies of services, which they map to the NCANDS services categories. Some known difficulties with reporting services data include:

- Children and families may receive services under more than one funding stream and may be counted more than once. Some programs count families, while

others count children.
- Prevention services are often provided by local community-based agencies, which may not be required to report on the number of clients they serve.
- Agencies that receive funding through different streams also may report to different agencies. CPS may have difficulty collecting data from all funders or all funded agencies.

Documentation:

Guidelines and procedures for submitting data to NCANDS are available to all authorized NCANDS website users. The guidelines includes a mapping form for each NCANDS field.

A User Guide designed for researchers using the NDACAN version of the NCANDS Child File dataset is publicly available at https://www.ndacan.cornell.edu/datasets/datasets-list-ncands-child-file.cfm.

Overall quality:

Overall, NCANDS data are considered to be of high quality by experts due in part to the fact that it serves as a major source for federal assessment of state performance. There are, however, notable variations in data quality by state and by topical area. A validation process occurs for all data submission files using code checking and inter- and cross-record rule checking to ensure compliance with the federal government's and NCANDS' data standards. The NCANDS Technical Team developed a cloud-based online data collection and validation system called NCANDS website, which integrates the data collection and validation procedures. The primary output from the validation process is a state dataset that is ready for analyses in conjunction with the child maltreatment data from all other states. Each year, the NCANDS Team generates multiyear, state-by-state data quality and completeness reports after validation to evaluate improvements and to identify any new issues. The reports are loaded into the Data Quality Matrix, which provides a comprehensive list of significant errors and recommended data improvements.

Program Scope and Budget

Size of population served annually:

During FY 2017, approximately 3,501,000 children received an investigation or alternative response. Of those, 674,000 were determined as victims of maltreatment. Source: https://www.acf.hhs.gov/cb/resource/child-maltreatment-2017.

Federal funding for programs encompassed by data:

$29.9 billion per year in federal, state, and local expenditures in SFY 2016, including $7.5 billion in Title IV-E funds. Source: Child Trends, available at https://www.childtrends.org/wp-content/uploads/2018/12/CWFSReportSFY2016_ChildTrends_December2018.pdf

National Directory of New Hires (NDNH)

Website

https://www.acf.hhs.gov/css/resource/a-guide-to-the-national-directory-of-new-hires

Funding Agency and Data Ownership

OCSE, ACF/HHS

Data Gathering Agency/Contractor

OCSE, ACF/HHS

Overview

The NDNH is a national repository of employment, UI, and quarterly wage (QW) information. The data residing in the NDNH includes records from the State Directory of New Hires, QW and UI data from state workforce agencies (SWAs), and new hire and QW data from federal agencies.

Periodicity

Continuous. New hire information is submitted to the state within 20 days of hire, and within another eight business days to NDNH. NDNH is updated daily. SWAs transmit QW data to the NDNH within four months of the end of a calendar quarter. Federal agencies transmit QW data to the NDNH no later than one month after the end of a calendar quarter. SWAs transmit UI data within one month of the end of a calendar quarter.

Data Content and Structure

- New Hire File: Employee name, employee SSN, employee address, employer name, Federal Employer Identification Number (FEIN), employer address, and date of hire.
- QW File: Employee name (collected by most states), employee SSN, employee wage amount, reporting period, employer name, FEIN, employer address, and employer optional address.
- UI File: Claimant name, claimant SSN, claimant address, claimant benefit amount, and reporting period.

Major Publications

A Guide to the National Directory of New Hires: https://www.acf.hhs.gov/css/resource/a-guide-to-the-national-directory-of-new-hires

NDNH Guide for Data Submission: https://www.acf.hhs.gov/css/resource/ndnh-guide-for-data-submission

2016 Preliminary Report (see page 101):
https://www.acf.hhs.gov/sites/default/files/programs/css/2016_preliminary_report.pdf

Available Datasets

Public-use aggregate:

None.

Public-use micro:

The Secretary may provide access to the NDNH data, without personal identifiers, for research purposes found by the Secretary to be likely to contribute to achieving the purposes of Part A or Part D, which refers to TANF and Child Support Enforcement. See https://www.ssa.gov/OP_Home/ssact/title04/0401.htm.

Restricted-use micro:

The NDNH is a restricted use dataset as it includes PII. OCSE cannot disclose NDNH information if the law does not specifically authorize an agency to receive specified NDNH information and the information or comparison being requested does not meet the purposes stated in the statutory authority.

In-house micro:

Access to the NDNH is granted to other federal agencies by Congress and authorized in Section 453 of the Social Security Act. Federal law provides that a state or federal agency that receives NDNH information must reimburse OCSE for the costs of obtaining, verifying, maintaining, and comparing the information at rates OCSE determines to be reasonable.

Statutory and Regulatory Restrictions

Title IV-D of the Social Security Act, which governs the NDNH, specifies the persons or entities authorized to request NDNH information and the purposes for which the information may be used. 42 U.S.C. 653(i)(3) and (j)(3-11). Statutory authority is required to receive NDNH information. OCSE cannot disclose NDNH information if the law does not specifically authorize an agency to receive specified NDNH information and the information or comparison being requested does not meet the purposes stated in the statutory authority. 42 U.S.C. 653(j)(5) provides NDNH information without personal identifiers to conduct research found by the HHS Secretary to be likely to contribute to achieving the purposes of Part A or Part D of the Social Security Act. See https://www.ssa.gov/OP_Home/ssact/title04/0453.htm.

Feasibility to Link to Other Data

Personally identifiable information:

The NDNH database contains three files: New Hire, QW, and UI. Data elements collected are listed below.

- New Hire File: Employee name, employee SSN, employee address, employer name, FEIN, employer address, and date of hire.
- QW File: Employee name (if collected by state), employee SSN, employee wage amount, reporting period (calendar quarter in which wages were paid), employer name, FEIN, employer address, and employer optional address.
- UI File: Claimant name, claimant SSN, claimant address, claimant benefit amount (gross amount before any deductions), and reporting period (calendar quarter in which the UI claim was filed).

Geographic and institutional identifiers:

- New Hire File: Employee address, employer name, FEIN, and employer address.
- QW File: Employer name, FEIN, employer address, and employer optional address.
- UI File: Claimant address.

Data it has been linked to in the past:

At least five federal agencies representing 20 data matches with OCSE for research purposes.

Evaluations that have done such matches include: Maternal and Infant Home Visiting Program Evaluation (MiHOPE), National Evaluation of Welfare-to-Work Strategies (NEWWS), Youthbuild, Health Profession Opportunity Grants (HPOG) Program Evaluation, Pathways for Advancing Careers and Education (PACE), Subsidized and Transitional Employment Demonstration (STED), Parents and Children Together (PACT) Evaluation, the Child Support Noncustodial Parent Employment Demonstration (CSPED), Enhanced Transitional Jobs Demonstration (ETJD), and the Workforce Investment Act (WIA) Gold Standard Evaluation.

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

Federal agencies administering programs, such as SNAP, and social science researchers are interested in matching with the employment and earnings data in the NDNH to assess the effectiveness of programs, interventions, and other efforts.

Data Quality

NDNH data are generally regarded as high-quality administrative data. OCSE assesses the collection of the data during annual quality assurance audits in select states.

Validity and reliability:

The interpretation of the variables and how users enter the data is uniform. OCSE assesses the collection of the data during annual quality assurance audits in select states.

Nonresponse and coverage:

Federal and nonfederal employers are required by law to submit new hire information to NDNH and to their state within 20 days of hire, thus covering the intended population. Currently, independent contractors are not required to report employment and wage information to the state directories of new hires. When the data are received by NDNH, the system conducts edits and the name and SSN combinations also go through a verification process.

Documentation:

Information on the HHS website includes: NDNH history; statutory authority; comprehensive guides containing information about available data and how it is collected, how frequently it is collected, and the sources of the information; a list of the authorized federal and state agencies and the specific purposes for which each agency may use the NDNH; and various reports demonstrating the benefits of using the NDNH. See https://www.acf.hhs.gov/css/resource/a-guide-to-the-national-directory-of-new-hires.

Overall quality:

Data quality for the NDNH is generally considered to be high by subject matter experts. Information in the NDNH is the most current and accurate employment and wage information that is available.

Program Scope and Budget

Size of population served annually:

Total number of records in FY 2016: 741,686,791.

Source:
https://www.acf.hhs.gov/sites/default/files/programs/css/fy_2016_annual_report.pdf, Table 97.

Federal funding for programs encompassed by data:

Total expenditures for the child support program in FY 2016 were $5.7 billion ($3.6 billion in federal funds and $2.3 billion in state funds). Source: https://www.acf.hhs.gov/sites/default/files/programs/css/fy_2016_annual_report.pdf.

National Domestic Violence Hotline Advocate Caller Application Database

Website

None

Funding Agency and Data Ownership

Family Violence Prevention Services Act (FVPSA) Program, FYSB/ACYF/ACF/HHS

Data Gathering Agency/Contractor

National Domestic Violence Hotline

Overview

The Advocate Caller Application database includes information about each contact to the National Domestic Violence Hotline (The Hotline) or loveisrespect (LIR) helpline, made by telephone, chat, text, e-mail, or social media. This information is entered into the database manually by advocates at the time of contact. It is primarily used for service provision and operational purposes. It does not include any PII.

Periodicity

Continuous. Data are collected on every contact made to The Hotline or LIR, which operate 24 hours a day, seven days a week

Data Content and Structure

The Advocate Caller Application database includes demographic information about the person who called, chatted, texted, etc., and his/her situation (e.g., type of abuse), and information about what happened during the call, chat, or text (e.g., topics discussed, services provided, etc.). It also includes information about caller needs and reported barriers to receiving services.

Major Publications

Accomplishments of the Domestic Violence Hotline, Online Connections, and Text (ADVHOCaT) brief report: https://www.acf.hhs.gov/opre/resource/accomplishments-domestic-violence-hotline-online-connections-text-initial-findings-next-steps

FVPSA Report to Congress:
https://www.acf.hhs.gov/sites/default/files/fysb/fvpsa_report_to_congress_09_10.pdf

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

The Hotline and LIR have made these data available to OPRE contractors for the purpose of the ADVHOCaT and Summative ADEPT Formal Evaluation of Classroom-Based Teachers (SAFE-T) evaluations.

In-house micro:

None.

Statutory and Regulatory Restrictions

FVPSA includes statutory restrictions on collecting PII. See Section 10406(c) (5) of FVPSA statute at https://www.gpo.gov/fdsys/pkg/USCODE-2010-title42/html/USCODE-2010-title42-chap110.htm.

Feasibility to Link to Other Data

Personally identifiable information:

None.

Geographic and institutional identifiers:

City and state.

Data it has been linked to in the past:

None.

Example research products using linked data:

None.

Data Quality

Validity and reliability:

The data in the Advocate Caller Application database are often incomplete for each contact. The advocate does not directly ask the contactor about each data element, so the data entered is dependent on what the contactor chooses to share. Also, the advocate is providing services and often dealing with crises at the same time as entering data, which limits the consistency and accuracy of the data.

Nonresponse and coverage

See above.

Documentation:

None

Overall quality:

The overall data quality of the Advocate Caller Application database is considered moderate by subject matter experts, with notable issues as outlined above.

Program Scope and Budget

Size of population served annually:

Between January 2014 and December 2015, The Hotline and LIR received 503,620 contacts. (Because they do not collect PII, there is no way of knowing if those contacts are from unique people; they include repeat contactors.)

Federal funding for programs encompassed by data:

$8.25 million in FY 2016 for The Hotline services.

National Youth in Transition Database (NYTD)

Website

https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/nytd

Funding Agency and Data Ownership

Children's Bureau, ACYF/ACF/HHS

Data Gathering Agency/Contractor

Children's Bureau is the data-gathering agency. Under contract with the Children's Bureau, ICF International developed the portal for states to submit NYTD data.

Overview

Information on NYTD Services:
The regulation requires that States report to ACF information on all youth and young adults currently or formerly in foster care who received independent living services (the "served population") paid for or provided by the state agency that administers the John Chafee Foster Care Independence Program (CFCIP). Services are reported under 14 broad categories: independent living needs assessment; academic support; postsecondary educational support; career preparation; employment programs or vocational training; budget and financial management; housing education and home management training; health education and risk prevention; family support and healthy marriage education; mentoring; supervised independent living; education financial assistance; room and board financial assistance; and other financial assistance.

Information on NYTD Outcomes:
States will survey youth regarding six outcomes: financial self-sufficiency, experience with homelessness, educational attainment, positive connections with adults, high-risk behavior, and access to health insurance.

Survey Requirements:
States are to collect outcomes information by conducting a baseline survey of all youth in foster care within 45 days of their turning age 17 while in care (the "baseline population"). States will follow the youth who completed the survey at age 17 and conduct a new outcome survey on or around the youth's 19th birthday; and again on or around the youth's 21st birthday, also referred to as the follow-up population. States will collect outcomes information on these older youth at ages 19 or 21 regardless of their foster care status. Some States may opt to follow a random sample of the 17-year-olds who participated in the outcomes survey so that they can follow a smaller group of

youth at ages 19 and 21. All states will collect and report outcome information on a new baseline population cohort every three years, beginning with the cohort of youth turning age 17 in FY 2011. More details can be found at https://www.acf.hhs.gov/cb/resource/about-nytd?page=2.

Periodicity

States are required to submit NYTD data semi-annually to the Children's Bureau. New cohorts are started and reported every three years. The first cohort began in FY 2011.

Data Content and Structure

Data elements reported in NYTD include demographics such as:

- Date of birth
- Race and ethnicity
- Sex
- Tribal membership
- Foster care status
- Educational level

They also report on services, such as:

- Academic support
- Career preparation
- Budgeting
- Mentoring
- Health education
- Financial assistance

And outcomes, including:

- Financial self-sufficiency
- Educational attainment
- Homelessness
- High-risk behaviors
- Connections with adults
- Access to health insurance

States may add additional questions to their own surveys, though they do not report those data to CB.

Major Publications

NYTD data briefs are published on the Children's Bureau's website at https://www.acf.hhs.gov/cb/resource/data-briefs.

Available Datasets

Public-use aggregate:

There are no public use data files, but data briefs with high-level findings are available. Tabular data can be made available upon request from NDACAN at http://www.ndacan.acf.hhs.gov. Additional aggregate tables are published bi-annually in The Green Book. The NYTD Portal is available to state administrators and designated staff (e.g., independent living coordinators, data staff), and they should soon have the functionality to run a "static" report through the data portal. States can't do analyses through the Portal and the report feature allows states to choose reporting years and populations but with limited functionality. Of note, the Portal was not designed for analyses as states use their own administrative systems for that. Currently, states can put in a request to the Children's Bureau for a custom run.

Public-use micro:

None (microdata files are available to researchers only, see below).

Restricted-use micro:

NDACAN makes NYTD data available to researchers upon request. NYTD data are restricted access files, as only individuals with a research-related affiliation with an institution may request data, unless NDACAN grants an exception. Requestors must provide detailed contact information and sign a Terms of Use agreement. Undergraduate students may serve as investigators if a faculty advisor co-signs the agreement. They must promise confidentiality and never make any intentional identification of a research subject, or unauthorized disclosure of confidential information, and the research data must be used solely for research or statistical purposes. In addition, the data have been masked by the staff at NDACAN prior to distribution. Dates of birth have been rounded to the 15th of the month, and counties with fewer than 1,000 children have been recoded to mitigate risk of disclosure. All unique identifiers have been encrypted, and there are no PII. A user's guide for the Outcomes file can be found at http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset202UsersGuide.pdf, and a user's guide for the Services file may be found at http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset222UsersGuide.pdf.

In-house micro:

In-house datasets contain all state and county identifiers, as well as original values for measures that are rounded, suppressed, or perturbed in the datasets distributed by the NDACAN. Staff may produce estimates requested by other federal agencies or Congress on request (and subject to resource limitations

Statutory and Regulatory Restrictions

Public Law 106-169 established the John H. Chafee Foster Care Independence

Program (CFCIP) at section 477 of the Social Security Act, providing states with flexible funding to carry out programs that assist youth in making the transition from foster care to self-sufficiency. The law also requires ACF to develop a data collection system to track the independent living services states provide to youth and develop outcome measures that may be used to assess states' performance in operating their independent living programs. The law requires ACF to impose a penalty of between 1 percent and 5 percent of the state's annual allotment on any state that fails to comply with the reporting requirements.

To meet the law's mandate, ACF published a proposed rule in the Federal Register on July 14, 2006, and a final rule on February 26, 2008. The regulation establishes NYTD and requires that states engage in two data collection activities. First, states are to collect information on each youth who receives independent living services paid for or provided by the state agency that administers the CFCIP. Second, states are to collect demographic and outcome information on certain youth in foster care whom the state will follow over time to collect additional outcome information. This information will allow ACF to track which independent living services states provide and assess the collective outcomes of youth. Pursuant to the regulation, states began collecting data for NYTD on October 1, 2010, and report data to ACF semiannually. The first submission of data to ACF was due May 15, 2011.

NYTD guidance is available at https://www.acf.hhs.gov/cb/resource/nytd-guidance.

Feasibility to Link to Other Data

Personally identifiable information:

For every youth reported to NYTD, a state must use an encrypted identification number that is the same as the identifier used to report information on the young person to AFCARS. This enables analysis on the information related to a youth's foster care experiences reported to AFCARS along with their outcomes and/or services information reported to NYTD. A state must also report to NYTD the youth's sex, race, ethnicity, date of birth and foster care status. When a state reports on independent living services, it must identify the local agency responsible for the youth, whether the youth is a member of a federally recognized Indian tribe, the youth's educational level, the youth's receipt of special education, and whether the youth has been adjudicated delinquent. (https://www.acf.hhs.gov/cb/resource/about-nytd?page=all).

Encrypted identifiers are included that allow for linking over time and with AFCARS and NCANDS records only. Exact date of birth and other date information that could be used for identification are collected, though only perturbed versions are made available in datasets distributed by NDACAN.

Geographic and institutional identifiers:

State and county identifiers are available in the datasets available through NDACAN. The county represents the county with responsibility for the care of the child, rather than the child's residence. County identifiers are suppressed if there are fewer than 1,000 records reported. NDACAN is examining alternatives for making county identifiers available through a restricted access virtual data center. In-house datasets contain all county identifiers, but access is limited to federal staff.

Data it has been linked to in the past:

The variable RecNumbr is an encrypted version of the child's unique identifier used by the state agency. The ID may go by different names in the various linkable files. These are:

- NYTD Outcomes File: RecNumbr
- NYTD Services File: RecNumbr
- AFCARS Foster Care File: RecNumbr
- AFCARS Adoption File: RecNum
- NCANDS Child File: AFCARSID

Generally speaking, these IDs are the same and can be linked across NCANDS, AFCARS foster care files, and NYTD files. The only ID that is different and will not find a match is the AFCARS Adoption File RecNum, which is done intentionally to prevent linkage, to preserve confidentiality. All of the IDs are encrypted, and are encrypted in the same way for all these datasets, so it serves as an indicator of the same child across datasets and across years. These commonalities are generally reliable, but are not applicable to all states in all years. Contact NDACAN Support for further information regarding which states can be linked across which years. See http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset202UsersGuide.pdf.

Example research products using linked data:

Shpiegel, S., Cascardi, M., & Dineen, M. (2016). A social ecology analysis of childbirth among females emancipating from foster care. Journal of Adolescent Health. https://doi.org/10.1016/j.jadohealth.2016.09.013

Shpiegel, S., & Ocasio, K. (2015). Functioning patterns among older adolescents in foster care: Results from a cluster analysis. Children and Youth Services Review, 58, 227–235. https://doi.org/10.1016/j.childyouth.2015.09.024

Okpych, N. J. (2015). Receipt of independent living services among older youth in foster care: An analysis of national data from the U.S. Children and Youth Services Review, 51, 74–86. https://doi.org/10.1016/j.childyouth.2015.01.021

Potential Research Areas to Inform Policy

Data can be used to study outcomes for youth transitioning from foster care at age 21, including employment, public benefits usage, and criminal justice indicators. Linking NYTD data to benefits data such as SNAP or TANF, Medicaid records, and/or court data, would be useful to learn more about youth cohorts as they move farther into adulthood.

Data Quality

Validity and reliability:

NYTD has established data reporting standards and penalties to grantees for not meeting those standards. For a state's data to comply with the NYTD standards, it must:

- Meet file format requirements and contain error-free information for certain standard demographic information;
- Contain information that is 90 percent error-free for other data elements (i.e., be free of missing information, internally inconsistent responses and invalid information);
- Provide full or partial outcome survey information on all 19- and 21-year-olds in the follow-up population or sample, or indicate why the survey information was not obtainable (i.e., because the youth is incapacitated or deceased); and
- Garner the participation in the outcomes survey of at least 60 percent of 19- or 21-year-old youth in the follow-up population who are no longer in foster care. For youth who are still in foster care as described at 45 CFR 1355.20 at ages 19 or 21, the state must achieve an outcome survey participation rate of at least 80 percent.

If the state does not meet these standards, the state will have an opportunity to transmit corrected data by the end of the subsequent report period. If the corrected data still do not meet the standards, the state will be penalized between one percent and five percent of their annual CFCIP allotment for each reporting period, depending on the standard that was not achieved. Education and Training Voucher funds (also authorized in Section 477 of the Social Security Act) will not be subject to a penalty if a state does not comply with these standards. More details can be found at https://www.acf.hhs.gov/cb/resource/about-nytd?page=3.

Nonresponse and coverage:

NYTD data is reported for two populations. The Outcomes data are the results of a survey. In the baseline population of 17-year-olds in foster care in FY 2011, 53 percent were surveyed. In the follow-up at age 19, approximately 67 percent were surveyed, and at 21 years of age in FY 2015, 60 percent were surveyed. The Services file should have complete coverage of the intended population, because states are required to report on all services provided under CFCIP. (A State is to submit a single data file to

ACF semiannually comprised of data on all youth in the served population, and if applicable, youth in the baseline or follow-up population [45 CFR 13656.82 and 1356.83(a) through (e)].). In a year in which data collection is not required on the baseline or follow-up populations (i.e., FY 2012 when the youth in the initial baseline population are 18 years of age), a state must still submit to ACF a data file composed of youth of any age who are in the served population for a report period. See https://www.childwelfare.gov/cb/research-data-technology/reporting-systems/nytd/faq/reporting-populations-and-reporting-requirements.

Documentation:

See user guides available at NDACAN: http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset202UsersGuide.pdf and http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset203UsersGuide.pdf.

Overall quality:

States continue to provide increasingly reliable data to NYTD, and most subject matter experts consider the general quality of the data to be high. States are able to run data quality checks before submitting their data to NYTD using a tool called the NYTD Data Review Utility (NDRU), available at https://nytd.acf.hhs.gov/ndru/.

Program Scope and Budget

Size of population served annually:

On September 30, 2017, there were 37,779 youth ages 17 and older in foster care, about 9 percent of all children in foster care. Source: https://www.acf.hhs.gov/sites/default/files/cb/afcarsreport25.pdf.

Federal funding for programs encompassed by data:

In FY 2019, there was $140 million in mandatory funding for the Chafee Foster Care Independence Program, and an additional $43 million in funding for the Education and Training Voucher Program (ETV). Source: https://fas.org/sgp/crs/misc/IF11070.pdf

OCSE Debtor File

Website

None.

Funding Agency and Data Ownership

OCSE, ACF/HHS

Data Gathering Agency/Contractor

OCSE

Overview

The OCSE Debtor File is a repository of persons who owe past-due child support for a case enforced by the child support program, referred to as debtors, and the amount of their past-due child support debts.

A single submission procedure is used for states submitting cases to the OCSE Debtor Master File for each of the following five remedies:

1) Federal income tax refund offset;
2) Federal administrative offset;
3) Denial of U.S. passports;
4) Multistate Financial Institution Data Match (MSFIDM); and
5) Federal insurance match and debt inquiry.

States are required to submit all cases that meet the criteria for federal income tax refund offset to OCSE for collection through the OCSE debtor file. In addition, states must have procedures in place to participate in the passport denial program and MSFIDM. Administrative offset and federal insurance match are optional programs. Cases are submitted to all five programs if they meet the programs' eligibility criteria and are not specifically excluded from that remedy by the state.

OCSE, through the Debtor File, provides information to the Department of Treasury for Federal Income Tax Refund Offset and Administrative Offset Programs, to the U.S. Department of State for the Passport Denial program, to multistate financial institutions for matching of financial accounts, and to insurance providers for matching insurance claim information.

Periodicity

State agencies submit data periodically, often in batches, to the OCSE Debtor File.

Data Content and Structure

The OCSE Debtor File collects from state child support enforcement agencies information pertaining to past-due child support owed by noncustodial parents, including PII such as the name and SSN of such individual; the amount of past-due child support owed by the individual and adjustments to such amount; information on each enforcement remedy being requested that is applicable to the individual to whom the record pertains, as indicated by a state child support agency; the amount of past-due support collected as a result of each remedy; and a history of updates by the state agency or agencies certifying the debtor.

Major Publications

OCSE FY 2016 Annual Report to Congress, page 8:
https://www.acf.hhs.gov/css/resource/fy-2016-annual-report-to-congress

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

The OCSE Debtor File is a restricted use dataset. States may submit inquiries, submit transactions, and perform other activities on their certified cases through various applications on the Child Support Portal. OCSE manages and maintains the Child Support Portal.

In-house micro:

OCSE exchanges information with Treasury’s Bureau of the Fiscal Service for federal income tax refund and administrative offset purposes. OCSE also matches with multistate financial institutions and insurers and returns that information to the states.

Statutory and Regulatory Restrictions

The data collection is mandatory and federal law requires collection of information for child support purposes and other authorized purposes. 42 USC Section 654(26) and 42 USC Section 653 (l) and (m).

Feasibility to Link to Other Data

Personally identifiable information:

PII in the OCSE Debtor File includes name, SSN, date of birth, taxpayer ID, current address, financial accounts information, place of birth, and child support arrearages and payment updates.

Geographic and institutional identifiers:

Data is restricted. Current address and date of birth are available in the OCSE Debtor File.

Data it has been linked to in the past:

This dataset has been linked for research purposes by OCSE. It was linked to other datasets in OCSE's Federal Parent Locator Service (FPLS).

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

Linking the OCSE Debtor File to Census data would enable policymakers to have possible demographic information on the debtor caseload, although there are caveats where the address could be out of date or inaccurate. This linkage could help the program better explore how to serve these cases and understand the nature of child support debt.

Data Quality

Generally, the OCSE Debtor File is considered a good administrative dataset. It can be compared with other administrative data in the FPLS and other federal systems. Name and SSN combinations are verified by the SSA.

Validity and reliability:

Data is continually added, deleted, and updated to ensure synchronization between states, OCSE, and their federal and private partners. OCSE conducts annual data reliability audits on state child support data which is the source of data for the OCSE Debtor File. States are also required annually to complete and return an Annual Certification Letter verifying and certifying that all of its debts are accurate and meet federal requirements.

The OCSE Debtor file may have limited out-of-date information for variables such as current address. This may impede the location of noncustodial parents and their assets, if any. OCSE requires states to submit updates, changes, and eliminations of past-due amounts to the OCSE Debtor File at least biweekly, though most states update weekly.

Documentation:

OCSE produces and maintains technical guides and program instructions for the OCSE Debtor File data and programs for states and its partners. Training and technical assistance is also provided, and state-specific liaisons respond to questions.

Overall quality:

Data quality of the OCSE debtor file is considered to be high by subject matter experts. However, the quality is contingent on what is provided to OCSE by states and its federal agency partners.

Program Scope and Budget

Size of population served annually:

The total number of debtors certified for the Federal Offset Program in FY 2016 was 7.6 million. Source: https://www.acf.hhs.gov/sites/default/files/programs/css/fy_2016_annual_report.pdf, Table 98.

Federal funding for programs encompassed by data:

Total expenditures for the child support program in FY 2016 were $5.7 billion ($3.4 billion in federal funds and $2.3 billion in state funds).

Office of Family Assistance (OFA) TANF Data Reporting System

Website

https://www.acf.hhs.gov/ofa/programs/tanf/data-reports

Funding Agency and Data Ownership

Office of Family Assistance (OFA), ACF/HHS

Data Gathering Agency/Contractor

Unissant

Overview

Each state is required to collect monthly and submit quarterly reports of individual-level data for four categories of persons: active TANF recipients, closed TANF cases, active participants in separate state programs (SSP), and closed SSP cases. States may opt to report information on all recipients or a stratified sample.

Periodicity

Monthly.

Data Content and Structure

Measures reported for current TANF recipients and SSP active cases; information on closed cases is somewhat more abbreviated.

- Demographics (family type, number of members, race, ethnicity, gender, parent status, relationship to household head, education level, citizenship/alienage)
- Geography (state, county, ZIP)
- Benefits/cash received (subsidized housing, child care, job training, education, community service, job search, number of months received)
- Sanctions
- PII (SSN, date of birth)
- Employment (experience, eligibility, status, public/private)
- Income (amount, type, source)

Major Publications

https://www.acf.hhs.gov/ofa/programs/tanf/data-reports

Available Datasets

Public-use aggregate:

TANF Caseload Data (annual), and TANF Work Participation Rates (annual): https://www.acf.hhs.gov/ofa/programs/tanf/data-reports.

Public-use micro:

None.

Restricted-use micro:

Datasets are housed at several data resource centers set up specifically to handle sensitive data. Access to the data are highly restricted to a limited number of researchers. CARRA (Census) is one location (https://www.census.gov/about/adrm/linkage.html), and the other is the Family Self-Sufficiency Data Center (https://harris.uchicago.edu/research-impact/centers/family-self-sufficiency-data-center).

In-house micro:

The in-house dataset is available only to the contractor and select OFA staff.

Statutory and Regulatory Restrictions

The data may be shared without consent pursuant to 5 U.S.C. 552a(b)(3), in accordance with routine use disclosures 1 (disclosure of identifiable data for research) and 7 (disclosure to contractors, grantees, and others) published for System No. 09-80-0375, last published at 80 Fed. Reg. 17894, 17903-05 (April 2, 2015).

Feasibility to Link to Other Data

Personally identifiable information:

SSN, date of birth, and case number.

Geographic and institutional identifiers:

State, county, and ZIP code.

Data it has been linked to in the past:

Current Population Survey, SNAP, Medicaid, child care, UI data, child welfare, birth records, and tax records.

For additional details, see “A Bibliography of Studies Using Temporary Assistance for Needy Families (TANF) Linked Administrative Data.” Paul Johnson, Jim Kaminski, Molly

Scott, and Anne Whitesell. 2012. Urban Institute. Available at https://www.urban.org/sites/default/files/publication/25261/412540-A-Bibliography-of-Studies-Using-Temporary-Assistance-for-Needy-Families-TANF-Linked-Administrative-Data.PDF.

Example research products using linked data:

For an annotated bibliography of over 100 research articles using linked TANF data, see "A Bibliography of Studies Using Temporary Assistance for Needy Families (TANF) Linked Administrative Data." Paul Johnson, Jim Kaminski, Molly Scott, and Anne Whitesell. 2012. Urban Institute. Available at https://www.urban.org/sites/default/files/publication/25261/412540-A-Bibliography-of-Studies-Using-Temporary-Assistance-for-Needy-Families-TANF-Linked-Administrative-Data.PDF.

Potential Research Areas to Inform Policy

Family formation and stability; child support and noncustodial parent involvement; and employment dynamics.

Data Quality

Validity and reliability:

The Family Self-Sufficiency Data Center is currently conducting data quality analyses to assess the data's suitability for research.

Nonresponse and coverage:

States have the option of handing in data on the entire universe of TANF recipients, or a representative sample. Currently 21 states hand in sample data, and the remainder (30) hand in universe data.

Documentation:

No codebooks are publicly available at present, though they may be available on a restricted basis through CARRA (Census Bureau) (https://www.census.gov/about/adrm/linkage.html).

Overall quality:

The Family Self-Sufficiency Data Center is currently conducting data quality analyses of TANF data from 2012-2015 for OFA, OPRE, and ACF/HHS to assess its suitability for use in research.

Program Scope and Budget

Size of population served annually:

In FY 2017, there were over 3.5 million recipients of TANF and SSP combined.

Federal funding for programs encompassed by data:

About $17 billion in TANF block grant funds in FY 2013. The TANF data focus only on cash assistance, which represents about a quarter of the block grant.

Refugee Arrivals Data System (RADS)

Website

Federal Register Notice for ORR Internet Refugee Arrivals Data System: System of Records Number 09-80-0325, https://www.federalregister.gov/documents/2016/07/18/2016-16812/privacy-act-of-1974-system-of-records-notice

Funding Agency and Data Ownership

Office of Refugee Resettlement (ORR)/ACF/HHS funds the RADS data system for storage and analysis of the data. Data on refugees are owned by U.S. Department of State.

Data Gathering Agency/Contractor

Arrivals and 90-day update data for entering refugees comes from U.S. Department of State's Worldwide Refugee Admissions Processing System (WRAPS). Selected WRAPS fields are shared with ORR for inclusion in RADS, but the interagency agreement states ownership remains with State. In ORR, RADS is processed and administered by a contractor, GDIT.

Overview

The RADS database draws on WRAPS for information on all refugee arrivals at entry and 90 days after arrival for refugees 16 years old and older. Resettlement agencies are charged by the U.S. Department of State to gather information during the reception and placement period and enter it into WRAPS. These data include baseline information about all individual refugees arriving in the country, as well as follow-up at the 90-day mark. These data are transferred to ORR and stored in the RADS database.

Once at ORR, WRAPS records are supplemented with records on other ORR-served populations (parolees, asylees, Cuban or Haitian Entrants, Amerasians, Iraqi or Afghani Special Immigrant Visa holders, or victims of trafficking). According to the Federal Register Notice, "Records are used by HHS/ACF/ORR to generate data needed to allocate funds for Formula Social Services and Targeted Assistance grants according to statutory formulas; extract samples for the Annual Survey of Refugees (ASR); and support other budget and grant requirements and data requests from within and outside ORR. This system of records does not collect new information but consolidates information on eligible populations obtained from other agencies."

Periodicity

Continuous.

Data Content and Structure

Includes a range of individual-level information on characteristics on refugees, including arrival date, nationality code, date of birth, languages spoken, and resettlement location (address and phone number, if available). After 90 days in the U.S., resettlement agencies provide updated information including updated contact information and program participation.

Major Publications

https://www.acf.hhs.gov/orr/resource/annual-orr-reports-to-congress

Available Datasets

Public-use aggregate:

To view refugee arrival data sorted by country of origin and state of initial resettlement in the United States, see report from Refugee Processing Center website: https://ireports.wrapsnet.org/

Public-use micro:

None.

Restricted-use micro:

The contractor for the ASR has access to the data for the purposes of pulling the survey sample. The contractor only receives the variables necessary to construct the sampling frame.

In-house micro:

RADS is a Privacy Act System of Records (System of Records # 09-80-0325). All fields are available in the in-house version, but access is limited to those who have an approved reason to use the data, e.g., required reporting, budget-related tasks, etc. See “Statutory Authority” below for more specifics on routine uses associated with this SOR. ORR does not have formal mechanisms for cost sharing.

Statutory and Regulatory Restrictions

The approved routine uses of RADS are: (1) disclosure for law enforcement purpose;

(2) disclosure for private relief legislation; (3) disclosure to Congressional Office on behalf of requesting individuals; (4) disclosure to Department of Justice or in proceedings; (5) disclosure to the National Archives; (6) disclosure to contractor; (7) disclosure for administrative claim, complaint, and appeal; (8) disclosure in connection with litigation; (9) disclosure incident to requesting information; (10) disclosure in the event of a security breach; and (11) disclosure for cybersecurity monitoring. The contractor (6) provision is broad: "Information may be disclosed to a contractor performing or working on a contract for HHS and who has a need to have access to the information in the performance of its duties or activities for HHS," but research is not an explicit routine use.

Feasibility to Link to Other Data

Personally identifiable information:

Name, alien number. Date of birth is recorded pre-entry and may be unreliable. First and 90-day residential addresses and telephone numbers are of variable quality, but ~75 percent of FY 2010-FY 2015 arrivals were able to be matched with U.S. Postal Service Change of Address database and TransUnion TLO batch lookup during the administration of the 2016 ASR.

Geographic and institutional identifiers:

In-house RADS includes arrival and 90-day state, ZIP code, and the name of the resettlement affiliate where case was assigned.

Data it has been linked to in the past:

The ASR has been matched to RADS data as part of the survey processing contract, and RADS fields used as validation for survey responses where applicable.

Because RADS data that rely on WRAPS records are owned by the State Department, any linking of these records requires approval by the State Department, rather than ORR. Currently WRAPS data are not approved to by linked to other data sources for research purposes.

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

Tabulations of RADS data are used to understand inflows and impacts of refugees, asylees, and other ORR-eligible populations, including by certain demographics (such as country of origin and age).

Linked RADS records with other data systems, such as state databases or Census records, could identify employment and integration-related outcomes of refugees, as well as study the implementation and effectiveness of different interventions.

Data Quality

Validity and reliability:

Many of the pre-arrival fields transferred into RADS from WRAPS are entered in field sites around the world. Date of birth is best-available, but typically set to 1/1/Year of Birth when exact date is unknown, especially for older arrivals. First residential address and phone number are often the address/phone number of the assigned resettlement agency, not of actual residence. Because many refugees change residences upon their arrival, address information, even at 90 days, is unstable.

Nonresponse and coverage:

RADS contains all data for all persons eligible for refugee-related services, which includes principal applicant and his/her family members (however these are not linked to a single case receiving specific services).

Documentation:

Data documentation – beyond names and labels – is only minimally available from ORR's RADS database contractor. As part of OPRE/ASPE internal work, some information about variable origin and definition has been compiled. As part of the ASR contract, the contractor has put together descriptive statistics of variable overlap (such as country of origin/nationality/citizenship – formal definitions of these fields are not available from ORR). Although ASPE has not formally spoken to the State Department about variable definitions, entry instructions, etc., it has a document that describes the interface between ORR and the RPC, which includes record formats, data formats, and conventions. More information may be available from the State Department for variables originating from WRAPS.

Overall quality:

Experts are confident that data in RADS originating from WRAPS represent a full universe of refugee entrants into the U.S. unable to comment on the quality of the data on other ORR-eligible populations (universe data on entrants and program data from Matching Grant and Unaccompanied Refugee Minor programs) contained in RADS.

Program Scope and Budget

Size of population served annually:

In Fiscal Year 2017, roughly 54,000 new refugees arrived in the U.S. that were eligible for ORR-funded benefits and services.

Federal funding for programs encompassed by data:

The initial resettlement period (between arrival and 90 days) is funded by the State Department, initially captured by WRAPS and then transferred to RADS. HHS funds the RADS system.

Runaway and Homeless Youth – Homeless Management Information System (RHY-HMIS)

Website

https://www.acf.hhs.gov/fysb/programs/runaway-homeless-youth (program site)

Funding Agency and Data Ownership

The Family Youth Services Bureau (FYSB), ACYF/ACF/HHS

Data Gathering Agency/Contractor

American Institutes for Research (AIR)

Overview

The data system is used to support national and regional planning and service provision for FYSB runaway and homeless youth programs. It consists of a repository (RhyPoint) that collects client-level information from grantees twice a year. It also consists of dashboards that visualize aggregate data collected by the repository. The dashboard features a grantee- and Federal Project Officer-specific component allowing users to view grantee-level information. The aggregate data are used for planning and policy, including ACF Performance Measures, the RHY Report to Congress, and other tools.

Periodicity

Twice per year.

Data Content and Structure

Data are collected on the following topics:

- Demographic information (DOB, race, ethnicity, gender, prior living situation, prior living situation length of time, disabling condition, sexual orientation), and veteran status
- Project start and exit dates
- Destination at exit
- Project completion status
- Safe and appropriate exits and permanent connections
- Client location
- Relationship to head of household (to collect information about unaccompanied

youth and their children or other individuals who present with the youth)

- Income, benefits, health insurance
- Special needs (including physical disability, chronic health condition, developmental disability, mental health issue, and substance abuse)
- Health status (general health, dental health, and mental health)
- Family critical issues
- Contacts and engagements established (for Street Outreach Program grantees only)
- Referral source
- Services and referrals provided (RHY service connections)
- Education and employment statuses
- Pregnancy status
- Former ward of foster care or juvenile justice
- Commercial sexual exploitation/sex trafficking
- Labor trafficking
- Aftercare provided

Source: https://www.hudexchange.info/resource/3824/hmis-data-dictionary/.

Major Publications

Biannual Report to Congress on RHY:
https://www.acf.hhs.gov/sites/default/files/fysb/rhy_report_to_congress_fy1213.pdf.

Public dashboard is under development.

Available Datasets

Public-use aggregate:

Data through 2014 are available in a searchable online format through https://extranet.acf.hhs.gov/rhymis/. More recent data are expected to be made available through an online dashboard. The public-facing dashboard will include aggregate data organized by nation, state, and region. (See details below.)

Public-use micro:

None.

Restricted-use micro:

The public-facing dashboard will contain aggregate data which can be filtered by characteristics (such as race, gender, age, and program type) and geography (nation, region, state). The restricted-use data use a suppression methodology to prevent users from ascertaining personal identity that by adding a random integer from -10 to 10, inclusive, to each row, to distort the totals slightly. By distorting the totals, users are prevented from using complementary cells to ascertaining small sample sizes and the identity of those samples. The dashboard also does not publish any results for a table with a sample size less than 50.

In addition to the public-facing dashboard, there will be a component of the dashboard available only to grantees and Federal project officers to see grantee-specific data. This will require a user account and a two-factor authentication. This set of data will not use any suppression methodology. It will allow grantees and Federal project officers to compare measures and outcomes for a particular grantee and compare those measures with the aggregates for the nation, ACF region, and state.

In-house micro:

In-house datasets exist for use by staff and selected contractors working for FYSB. This data includes the raw data uploaded from grantees into the repository. No other federal agencies currently have access to the raw data.

Statutory and Regulatory Restrictions

The RHY Act generally prohibits the sharing of identifying information about youth served by the program. 42 U.S.C. §§5712(b)(7) and 5731. In addition, the RHY Final Rule include the same prohibitions. 45 CFR §1351.19(b)(1). Data are submitted to FYSB without PII, though all records have a unique person identifier that allows for linking records across years within RHY-HMIS.

Feasibility to Link to Other Data

Personally identifiable information:

Prior to data collection, names and SSNs are hashed (or hidden) using the SHA-256 algorithm. Thus, the data contain hexadecimal strings representing the name and SSNs of the clients without actually disclosing their identity. The hashed values allow for the identification of duplicates within the dataset. In addition, the full, unhashed date of birth is collected.

Geographic and institutional identifiers:

The local grantee is identified, as is the local Continuum of Care (CoC) in which it operates. A CoC is an administrative entity that may span a county or city, a combination thereof, or even the entire state.

Data it has been linked to in the past:

RHY-HMIS has not been linked to any other administrative datasets for research purposes.

Potential Research Areas to Inform Policy

- What are the prevalence and characteristics of runaway and homeless youth served by the RHY programs (could be linked with HUD HMIS data for additional prevalence)?
- Where are the largest and smallest gaps in employment and educational achievement for runaway and homeless youth (could be linked with LEA-level data from the McKinney-Vento Program)?
- How do RHY programs affect outcomes for runaway and homeless youth with mental illness and/or substance abuse?
- How do service delivery and outcomes differ in rural versus non-rural areas?
- To what extent do levels of private and public funding affect outcomes for runaway and homeless youth?
- What are the service delivery patterns and outcomes for faith-based organizations that serve runaway and homeless youth?
- What are the prevalence, characteristics, and outcomes of trafficked youth and what areas have the highest prevalence?

Data Quality

Data quality varies by grantee and by measure. Data completeness and quality are tracked in three different ways: 1) the completeness rate of grantees who upload data during each biannual process; 2) a grantee-level data completeness and quality report following an upload; and 3) a frequency report indicating data completeness and quality for each measure for each upload period.

During each biannual submission period, over 95 percent of grantees usually submit their data, indicating that about 3 percent to 5 percent of grantees are not represented in the dataset. Grantee-level data completeness reports indicate an overall high level of data completeness and quality for the Basic Center Program (BCP) and Transitional Living Program (TLP)/Maternity Group Home (MGH) programs (over 90 percent). The data completeness is lower for Street Outreach (above 80 percent) given the difficult circumstances in which data are collected (on the street or other precarious locations).

The frequency reports for measures show that completeness and quality do vary greatly for each measure. Measure-specific completion rates range from less than 1 percent for measures on sexual exploitation and human trafficking to 83 percent for measures on destination at exit. Additional technical assistance efforts, including webinars, presentations, and written guidance in addition to one-on-one consultations by FYSB staff, contractors, and other experts, are contributing to higher data quality.

Validity and reliability:

There is a federal data standards manual and data dictionary intended to serve as the authoritative source of data collection. However, questions asked by grantees and the range of data reported reveal that interpretation is quite variable, particularly for some measures. There are a few forces driving variable interpretation:

1) The RHY-HMIS system has integrated old RHY-HMIS data standards collected by RHY grantees with HMIS data standards, traditionally used by HUD homeless programs. Familiarity with measures plays a role in understanding how to collect data. Completion rates are much higher for measures that have existed prior to the integration with HMIS than for measures introduced to RHY grantees after the integration process.
2) Because many measures are used by multiple federal agencies, interpretation of certain measures has changed over time from a narrow use that fit HUD programs to more variable interpretation that better fit other federal programs.
3) Data collection experience varies widely nationally, as there are over 26 HMIS vendors and the administration of each local HMIS lies in the hand of the HUD-funded CoC.

Focus groups have been conducted to better understand users' perspectives on data collection.

Nonresponse and coverage:

Given high completeness rates for the BCP and TLP and the relative ease in collecting data in residential programs (versus on the streets or other precarious locations), data collection for youth served in these residential programs more accurately report the population served. However, one important potential inclusion error (a subpopulation that should not be contained in the dataset) are youth who are currently involved in child welfare and juvenile justice systems and receive shelter services. Because BCPs often serve youth in these systems but fund them through other streams such as child welfare or juvenile justice, they may be inadvertently included in the RHY-HMIS dataset.

Given the challenges associated with data collection on the street or other locations, we expect that the dataset for the Street Outreach Program or BCP prevention services might suffer from an exclusion error – that is, youth who are served by these programs who are not captured in the dataset or whose data is incomplete due to challenges associated with data collection.

As mentioned in the data quality section above, the low completion rates for measures such as human trafficking mean the data are unreliable and the actual prevalence of the problem appears to be more significant among the population served.

Documentation:

HUD publishes a Data Standards Manual, Data Dictionary, and the RHY Program HMIS Manual. Together these document the standards, instructions, and specifications of data collection. In addition, the contractor, AIR, publishes a guide to assist users in uploading data to the national repository, a Frequently Asked Questions document with questions solicited from grantees, and a data quality report guide to help users understand data quality scores.

The RHY Program HMIS Manual is available at https://www.hudexchange.info/resource/4448/rhy-program-hmis-manual/.

HMIS Guides and Tools is available at https://www.hudexchange.info/programs/hmis/hmis-guides/#hmis-data-and-technical-standards.

Overall quality:

Efforts have been made to increase the quality and consistency of RHY-HMIS data over the last several years with the introduction of a new reporting system. As mentioned above, three major data quality issues persist: 1) the small subset of grantees (3-5 percent) that are not able to successfully upload data to the national repository; 2) the range of data quality scores for the grantees that do successfully upload data; and 3) the range of completeness frequencies for particular measures.

Program Scope and Budget

Size of population served annually:

Data from FY 2016 indicate that 20,663 youth were served in BCP shelters, 7,882 youth received prevention services through a BCP, 33,442 youth were served by SOP, and 5,435 youth stayed in TLPs. All numbers are unduplicated.

A recent study released be Chapin Hall's Voices of Youth Count project estimates 3.5 million youth ages 18-25 experienced some form or homeless in the past 12 months. For youth ages 13-17, 700,000 experienced some form of homelessness in the prior 12 months. See http://voicesofyouthcount.org/brief/national-estimates-of-youth-homelessness/.

Federal funding for programs encompassed by data:

FY 2017 funding was $119,121,000 ($101,980,000 for BCP, TLP/MGH and $17,141,000 for SOP).

ACF SURVEY DATA RESOURCES

Annual Survey of Refugees (ASR)

Website

None

Funding Agency and Data Ownership

ORR, ACF/HHS

Data Gathering Agency/Contractor

Urban Institute

Overview

Since the 1980s, ORR has conducted the ASR, collecting information on refugees' experiences during their first five years in the U.S. The ASR is the only national, scientifically collected source of data on refugees' progress towards self-sufficiency. ORR uses the ASR results alongside other information sources to fulfill its Congressionally mandated reporting requirements under the Refugee Act of 1980.

The ASR instrument is currently undergoing its first revision in over 20 years. Current informed consent does not explicitly provide for linkages with administrative data. There is interest in geocoding respondent locations to include community context variables at a future date, and to better understand secondary migration from initial placement and survey date. For the first time, 2016 ASR microdata are archived and publicly available at ICPSR (https://www.openicpsr.org/openicpsr/project/104642/version/V3/view;jsessionid=9CA98ECC44FD2C76183AF5446BE219AA). The analytic file contains no PII, and four-category Census Region is the lowest level of geography available.

Periodicity

Annual.

Data Content and Structure

- Demographics: Basic demographic characteristics like age, gender, ethnicity, and country of origin, as well as information about secondary migration
- Pre-resettlement: Work experience and educational background achieved before refugees' arrival in the U.S.
- Economic Status: Information about employment, wages, and use of public

benefits after arrival in the U.S.
- Human Capital: Measures of refugees' skills, knowledge, and experiences relevant to employment, as well as investments in training or education in English language and technical skills since arrival in the U.S.
- Civic and Social: Measures participation in the receiving community and pursuit of engagement and membership in society since arrival in the U.S.
- Barriers and Facilitators: Information about housing, health, and family structure.

Major Publications

https://www.acf.hhs.gov/orr/resource/annual-orr-reports-to-congress

Available Datasets

Public-use aggregate:

None.

Public-use micro:

Public use microdata archived at ICPSR, available to the public with free registration. See https://www.icpsr.umich.edu/icpsrweb/.

Restricted-use micro:

None.

In-house micro:

ASR data from 2009 to present is housed with the COR in ORR. Access may be available to federal staff interested in using the files for research purposes.

Statutory and Regulatory Restrictions

Items from ASR are used in congressionally mandated reporting on the adjustment of refugees during the first five years in the U.S., particularly on their employment, English-language acquisition, and benefits use.

Feasibility to Link to Other Data

Personally identifiable information:

Varies by year of survey administration.

Geographic and institutional identifiers:

Varies by year of survey administration.

Data it has been linked to in the past:

The sample for the ASR is pulled from RADS. The link between ASR and updated contact information for survey administration (including current address/phone) is not preserved on the analytic file.

Example research products using linked data:

None.

Potential Research Areas to Inform Policy

ASR data can be used for a wide range of policy questions, including:

- The relationship between refugee characteristics and integration outcomes.
- Refugee employment and economic self-sufficiency outcomes
- Trends over time of characteristics, integration, and employment of different refugee arrival cohorts.
- Health characteristics of refugees.

Data Quality

The survey is administered by a single call center, and is offered in 17 languages. There is no systematic understanding of the quality or cultural appropriateness of survey translations, though the contractor debriefed with call center staff following 2016 fielding period; useful information was gained and will be incorporated. There are several concerns with the validity and cognitive burden of current survey items, particularly about questions on the sources of public benefits. Improvements to these questions are currently being field tested.

Validity and reliability:

We are confident in the quality of interviewer training and quality control during the fielding period for the 2016 ASR and forward; survey is now administered via CATI. Prior to 2016, the survey was administered via paper and pencil and data entry was a separate step. Skip patterns seem to be logically forced during data entry, but overall quality of the survey operation prior to 2016 is unknown. Cognitive interviewing during field testing has provided further information about the reliability and validity of survey questions.

Nonresponse and coverage:

Prior to 2016, sampling issues (unintentional oversampling of large households) and compounding non-response over a five-year longitudinal cohort seriously threatened validity of cross-sectional estimates. The Urban Institute has re-weighted the 2015 data to correct for these issues as much as possible, and the revised 2015 files are held at ACF. Beginning in 2016, the survey moved to a fresh cross-section of arrivals in the most recent five years; cooperation rates contingent on contact were very high (>75 percent). Effective location of the highly mobile population remains a challenge to this survey. Population is intentionally limited by cost/benefit for instrument translation; the 2016 survey covers >78 percent of all entrants, while the rest (from languages <1 percent of arrivals) are intentionally excluded. Item non-response is not a significant problem.

Documentation:

Documentation for ASR data prior to 2015 is extremely limited. Beginning with the 2016 ASR, documentation is improved and very detailed. Documentation for 2016, the most recent year currently available, is available at ICPSR.

Overall quality:

Prior to 2015, data quality of the ASR is considered moderate by subject matter experts, and known weaknesses are detailed above. From 2016 and beyond, the data quality is considered high.

Program Scope and Budget

State/local datasets:

None.

Size of population served annually:

Around 85,000 refugees were admitted to the U.S. in FY 2016.

Federal funding for programs encompassed by data:

FY 2016 funding for HHS refugee resettlement programs was $708 million.

Early Head Start Family and Child Experiences Study (Baby FACES 2009)

Website

https://www.acf.hhs.gov/opre/research/project/early-head-start-family-and-child-experiences-study-baby-faces

Funding Agency and Data Ownership

ACF, HHS

Data Gathering Agency/Contractor

Mathematica Policy Research

Overview

The Early Head Start Family and Child Experiences Survey (Baby FACES 2009) is a descriptive study of Early Head Start programs designed to inform program planning, technical assistance, and research at the national level. Baby FACES 2009 had five main goals:

1. Provide descriptive information about Early Head Start services offered, their frequency and intensity;
2. Identify key characteristics of families currently served in Early Head Start;
3. Investigate how programs individualize services to meet family needs;
4. Learn how Early Head Start children and families are faring over time; and
5. Explore associations between the type and quality of Early Head Start services and child and family well-being.

Baby FACES 2009 used a longitudinal cohort design and a nationally representative sample of 89 programs. It took a census of children in two age cohorts and followed them through their time in Early Head Start. The Newborn Cohort includes pregnant mothers and newborn children who were up to 8 weeks old at the time they were enrolled in the study (194 families were in this group in the spring of 2009) and the 1-year-old Cohort includes children who were between 10 and 15 months of age at the time they were enrolled in the study (782 families were in this group in the spring of 2009). Data were collected annually in the spring of each year until children aged out of Early Head Start (at about age 3) or left the program. Data collection started in the spring of 2009 and ended for the 1-year-old Cohort in spring 2011 and for the Newborn Cohort in spring 2012.

Periodicity

A longitudinal design was carried out from 2009 to 2012. Data were collected annually, in the spring, for four consecutive years. The study was redesigned as a repeated cross-sectional study and fielded in the spring of 2018 (second cohort); a third cohort is planned for spring of 2020.

Data Content and Structure

Parent Interview:
This interview asked the person primarily responsible for the care of the study child about demographic characteristics, their service needs and use, and their well-being and that of the child. It also asked about the child's exposure to environmental health risks and environmental and routine supports for the child's growth and development. Parents were also asked to rate their child's development and behavior.

Direct Child Assessments:
The assessments included administration of the Preschool Language Scale-4 Auditory Comprehension subscale (PLS-4; Zimmerman et al. 2002), the Peabody Picture Vocabulary Scale–Fourth Edition (PPVT-4; Dunn and Dunn 2007), and measurement of height and weight.

Home Observation:
While in the home, the field assessor also observed the child's ability to focus on the tasks, the interactions between the child and parents, and the quality of the home environment using the Bayley Behavior Rating Scale (BRS) (Bayley 2006), the Home Observation for Measurement of the Environment (HOME) (Caldwell and Bradley 2003), and scales drawn from a study of neighborhoods in Chicago (Ross et al. 2008).

Parent-Child Interaction:
Children participated in semi-structured interaction activities that involve playing with two different sets of toys.

Assessor-Child Interaction:
The field assessors interacted with children following the Early Communication Indicator (ECI) protocol (Greenwood et al. 2006), that was later scored from videos. Staff (i.e., Home Visitor or Teacher) Interview: The child's home visitor or the child's teacher were interviewed to determine her demographic characteristics, tenure working for the program, and well-being, as well as training and education experiences provided by the program and the work environment. Home visitors and teachers of study children also completed child-specific ratings of the participation of families in the program, the behavior of children using the Behavior Problems Index (BPI; Zill and Peterson 1986), and the quality of their relationship with the parents of study children.

Classroom Quality Observation:
Observers rated classrooms with the CLASS-T (Pianta et al. 2010), which measures the quality of teacher-child interactions in center-based settings and includes two subscales: Engaged Support for Learning and Emotional and Behavioral Support.

Home-Visit Quality Observation:
Field assessors observed the home visitors who provided services to children in the study sample using the Home Visit Rating Scales-Adapted (HOVRS-A) (Roggman et al. 2009) and a form that assessed the content and characteristics of the visit.

Family Services Tracking:
Early Head Start home visitors and teachers of study children completed a weekly service tracking form that detailed the number of service experiences (home visits or days in care) study children were offered and the number received.

For additional information, see the reports at https://www.acf.hhs.gov/opre/research/project/early-head-start-family-and-child-experiences-study-baby-faces.

Major Publications

All reports can be found at https://www.acf.hhs.gov/opre/research/project/early-head-start-family-and-child-experiences-study-baby-faces.

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

The data are restricted from general dissemination. Users interested in obtaining these data must complete a Restricted Data Use Agreement form and specify the reasons for the request. An application for access can be completed at http://www.researchconnections.org/childcare/studies/36074.

In-house micro:

None.

Statutory and Regulatory Restrictions

Original Statutory Authority: Head Start Act of 2007, Sec. 649. RESEARCH, DEMONSTRATIONS, AND EVALUATION [42 U.S.C. 9844]

“The Secretary shall carry out a continuing program of research, demonstration, and evaluation activities, in order to—

a. foster continuous improvement in the quality of the Head Start programs under this subchapter and in their effectiveness in enabling participating children and their families to succeed in school and otherwise; and (B) use the Head Start programs to develop, test, and disseminate new ideas based on existing scientifically valid research, for addressing the needs of low-income preschool children (including children with disabilities, homeless children, children who have been abused or neglected, and children in foster care) and their families and communities (including demonstrations of innovative non-center-based program models such as home-based and mobile programs), and otherwise to further the purposes of this subchapter.”

(B)

Statutory and regulatory restrictions on data collection, privacy, and use of data involving Head Start/Early Head Start are detailed in the Head Start Act of 2007 (https://eclkc.ohs.acf.hhs.gov/policy/head-start-act) and the Head Start Program Performance Standards (https://eclkc.ohs.acf.hhs.gov/policy/45-cfr-chap-xiii). These laws refer to HIPAA and FERPA, and include various restrictions specific to HS/EHS. Specifically: PUBLIC LAW 110–134—DEC. 12, 2007, Improving Head Start for School Readiness Act of 2007.

‘‘(4) CONFIDENTIALITY.—

‘‘(A) IN GENERAL.—The Secretary, through regulation, shall ensure the confidentiality of any personally identifiable data, information, and records collected or maintained under this subchapter by the Secretary and any Head Start agency. Such regulations shall provide the policies, protections, and rights equivalent to those provided to a parent, student, or educational agency or institution under section 444 of the General Education Provisions Act (20 U.S.C. 1232g).

‘‘(B) PROHIBITION ON NATIONWIDE DATABASE.—Nothing in this subsection shall be construed to authorize the development of a nationwide database of personally identifiable data, information, or records on children resulting from the use of measures under this subsection.

General Education Provisions Act, Sec. 444:

"except when collection of PII is specifically authorized by Federal law, any data collected by such officials shall be protected in a manner which will not permit the personal identification of students and their parents by other than those officials, and such personally identifiable data shall be destroyed when no longer needed for such audit, evaluation, and enforcement of Federal legal requirements."

Feasibility to Link to Other Data

Personally identifiable information:

There is no PII available for linking purposes, and consent forms did not include provisions for making PII available for any purpose.

Geographic and institutional identifiers:

No geographic identifiers are available in the archived datasets.

Data it has been linked to in the past:

To our knowledge, dataset has not been linked to any other datasets.

Example research products using linked data:

None.

Data Quality

Validity and reliability:

Data elements appear to be reasonably reliable. DCIs include a number of standard or established scales and measures that have been widely used in the field and assessed for reliability/validity. Any new measures/items developed by the contractor were pretested. Data were collected by a cadre of well trained professional data collectors.

Reliability of scales and inter-rater reliability of observational measures are available in the technical reports available at https://www.acf.hhs.gov/opre/research/project/early-head-start-family-and-child-experiences-study-baby-faces and in the User Guide available at http://www.researchconnections.org/childcare/studies/36074.

Additional reports that examined psychometric evidence of measures included "Measuring Infant/Toddler Language Development: Lessons Learned About Assessment and Screening Tools," available at https://www.acf.hhs.gov/opre/resource/measuring-infant-toddler-language-development-lessons-learned-about-assessment-and-screening-tools, and "Observed Quality and Psychometric Properties of the CLASS-T in the Early Head Start Family and Child Experiences Survey," available at https://www.acf.hhs.gov/opre/resource/observed-quality-and-psychometric-properties-of-the-class-t-in-the-early-head-start-family-and-child-experiences-survey.

Nonresponse and coverage:

The study used sample weights accounting for the sample design and survey nonresponse, so findings pertain to the Early Head Start programs nationally. The sample included fairly high rates of attrition from the Early Head Start Program, which made participants ineligible for study participation. A little more than 60 percent of participants stayed with the same program for three years or more.

Documentation:

A User Guide is available to the public at http://www.researchconnections.org/childcare/studies/36074.

Overall quality:

Data quality of Baby FACES is considered high by subject matter experts.

Program Scope and Budget

State/local datasets:

None.

Size of population served annually:

As of 2015-2016, Early Head Start was funded to serve 147,519 slots. Total cumulative enrollment including both pregnant women and children was 205,564. Source: https://eclkc.ohs.acf.hhs.gov/report/early-head-start-services-snapshot-2015-2016.

Federal funding for programs encompassed by data:

In 2015, federal funding for Head Start (inclusive of Early Head Start and all support activities) was $8,285,544,370. Source: https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html.

Head Start Family and Child Experiences Study (FACES)

Website

https://www.acf.hhs.gov/opre/research/project/head-start-family-and-child-experiences-survey-faces

Funding Agency and Data Ownership

ACF, HHS

Data Gathering Agency/Contractor

Mathematica Policy Research

Overview

The Head Start Family and Child Experiences Survey (FACES) provides descriptive information on the characteristics, experiences and outcomes of Head Start children and families, as well as the characteristics of the Head Start programs that serve them. Children are sampled from Head Start programs and centers from across the 50 states and the District of Columbia.

FACES 1997-2009
Five FACES cohorts were fielded from 1997 through 2009 – FACES 1997, 2000, 2003, 2006 and 2009. Each cohort through 2009 includes a nationally representative sample of 3- to 4-year-old children entering Head Start for the first time in the fall of the program year, their families, Head Start teachers, classrooms, centers, and programs. Children are sampled from Head Start programs and centers from across the 50 states and the District of Columbia.

FACES 2014-2018
In 2011, ACF redesigned the study in an effort to prioritize timely and relevant research questions, explore innovative design options, update assessment tools and measures, and better meet the data needs of the Office of Head Start for ongoing management of the Head Start program. The newly redesigned FACES, implemented in fall 2014, features a new “Core Plus” study design. The “Core” study includes regular, ongoing data collection on a set of key indicators at the program, classroom, and child/family level. The “Plus” options include rotating modules or one-time, special studies that allow for greater flexibility to respond to new policy issues and questions in a timely manner.

Data collection for the FACES CORE is fielded every two years alternating between a “Classroom + Child Outcomes Core” and a “Classroom Core” data collection. Specifically, the Classrooms will be observed every two years, while Child Outcomes will be observed every four years. The Classroom + Child Outcomes Core draws data from a nationally representative sample of 3- to 4-year-old children and their families enrolled in 60 Head Start programs, and from teachers, classrooms, centers, and programs (in those 60 plus an additional 120 randomly selected Head Start programs). The Classroom Core includes data on teachers, classrooms, centers, and programs in 180 randomly selected Head Start programs.

Source: https://www.acf.hhs.gov/opre/research/project/head-start-family-and-child-experiences-survey-faces.

Periodicity

Six FACES cohorts have been fielded to date – FACES 1997, 2000, 2003, 2006, and 2014. A seventh, FACES 2019, is underway and expected to collect the first round of data in the fall of 2019. In FACES 2014 and 2019 children and families are followed over the course of one program year. In prior cohorts, children in the FACES 1997 were followed until the end of first grade, 2000-2006 through kindergarten, and 2014 for one Head Start year only. Prior to 2014, FACES cohorts sampled newly entering children and the classrooms and programs serving them; in FACES 2014, sampling represented all children – whether newly entering or returning – and their classrooms and programs.

Data Content and Structure

At each wave of data collection, FACES children are administered a one-on-one assessment of their development that includes important school readiness skills in language, literacy, and mathematics, as well as measurement of their height and weight (FACES 2006, 2009, and 2014 only). Parents are interviewed and asked questions about their child’s health and development, family life, and Head Start experience. Head Start teachers and kindergarten teachers provide information about the sampled children’s classroom experiences and their social skills and behaviors. Head Start program and center directors and education coordinators provide information on program policies and practices. Head Start staff and kindergarten teachers also provide information on their own educational backgrounds and credentials. Classrooms are observed to measure structural and process quality, including teacher-child interaction.

Major Publications

All reports can be found at https://www.acf.hhs.gov/opre/research/project/head-start-family-and-child-experiences-survey-faces.

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

The 2014 Head Start Family and Child Experiences Survey (FACES 2014) data and User Guide were released in February 2018. Previous waves of data are available through http://www.researchconnections.org and also require a Restricted Data Use Agreement form and an application for access. See http://www.researchconnections.org/childcare/studies/36643.

In-house micro:

None.

Statutory and Regulatory Restrictions

Original Statutory Authority: Head Start Act of 2007, Sec. 649. RESEARCH, DEMONSTRATIONS, AND EVALUATION [42 U.S.C. 9844]

“The Secretary shall carry out a continuing program of research, demonstration, and evaluation activities, in order to--

(A) foster continuous improvement in the quality of the Head Start programs under this subchapter and in their effectiveness in enabling participating children and their families to succeed in school and otherwise; and (B) use the Head Start programs to develop, test, and disseminate new ideas based on existing scientifically valid research, for addressing the needs of low-income preschool children (including children with disabilities, homeless children, children who have been abused or neglected, and children in foster care) and their families and communities (including demonstrations of innovative non-center-based program models such as home-based and mobile programs), and otherwise to further the purposes of this subchapter.”

Statutory and regulatory restrictions on data collection, privacy, and use of data involving Head Start/Early Head Start are detailed in the Head Start Act of 2007 (https://eclkc.ohs.acf.hhs.gov/policy/head-start-act) and the Head Start Program Performance Standards (https://eclkc.ohs.acf.hhs.gov/policy/45-cfr-chap-xiii). These laws refer to HIPAA and FERPA, and include various restrictions specific to HS/EHS.

Specifically:

PUBLIC LAW 110–134—DEC. 12, 2007, Improving Head Start for School Readiness Act of 2007.

‘‘(4) CONFIDENTIALITY.—

‘‘(A) IN GENERAL.—The Secretary, through regulation, shall ensure the confidentiality of any personally identifiable data, information, and records collected or maintained under this subchapter by the Secretary and any Head Start agency. Such regulations shall provide the policies, protections, and rights equivalent to those provided to a parent, student, or educational agency or institution under section 444 of the General Education Provisions Act (20 U.S.C. 1232g).

‘‘(B) PROHIBITION ON NATIONWIDE DATABASE.—Nothing in this subsection shall be construed to authorize the development of a nationwide database of personally identifiable data, information, or records on children resulting from the use of measures under this subsection.

General Education Provisions Act, Sec. 444:

“except when collection of PII is specifically authorized by Federal law, any data collected by such officials shall be protected in a manner which will not permit the personal identification of students and their parents by other than those officials, and such personally identifiable data shall be destroyed when no longer needed for such audit, evaluation, and enforcement of Federal legal requirements.”

Feasibility to Link to Other Data

Personally identifiable information:

There is no PII available for linking purposes and consent forms did not include provisions for making PII available for any purpose.

Geographic and institutional identifiers:

No geographic identifiers are available in the archived datasets.

Data it has been linked to in the past:

Dataset has not been linked to any other datasets.

Example research products using linked data:

None.

Data Quality

Validity and reliability:

Data elements appear to be reasonably reliable. DCIs include a number of standard or established scales and measures that have been widely used in the field and assessed for reliability/validity. Any new measures/items developed by the contractor were pretested. Data were collected by a cadre of well-trained professional data collectors.

Reliability of scales and inter-rater reliability of observational measures are available in the technical reports available at https://www.acf.hhs.gov/opre/research/project/head-start-family-and-child-experiences-survey-faces and in the User Guides available for each wave at www.researchconnections.org.

Additional information on measures of child outcomes and of classroom quality are available at https://www.acf.hhs.gov/sites/default/files/opre/data_tables_for_child_outcomes_and_classroom_quality_in_faces_2009.pdf.

Nonresponse and coverage:

The study achieved high rates of completion for each of the DCIs over time. The study used sample weights accounting for the sample design and survey nonresponse, so findings pertain to the Head Start programs nationally.

Documentation:

User Guides for each wave are available to the public (for FACES 2014, user guides were released February 2018) and are available at www.researchconnections.org.

Overall quality:

Data quality of FACES is considered high by subject matter experts.

Program Scope and Budget

State/local datasets:

None.

Size of population served annually:

As of 2015, Head Start (inclusive of Early Head Start) was funded to serve 944,581 slots. Source: https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/head-start-fact-sheet-fy-2015.pdf.

Federal funding for programs encompassed by data:

In 2015, federal funding for Head Start (inclusive of Early Head Start and all support activities) was $8,285,544,370. Source: https://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/2015-hs-program-factsheet.html.

National Incidence Study (NIS)

Website

https://www.childwelfare.gov/topics/systemwide/statistics/nis/

Funding Agency and Data Ownership

Principal funding from Children's Bureau and OPRE, ACF/HHS. The National Institute of Child Health and Human Development (NICHD) contributed funding for the NIS–4 Sentinel Definitions Survey (SDS). Data are owned by the Children's Bureau and OPRE.

Data Gathering Agency/Contractor

Westat, Inc.

Overview

Although the Children's Bureau collects annual state-level administrative data on official reports of child maltreatment, the NIS studies are designed to estimate more broadly the incidence of child maltreatment in the U.S. by including both cases that are reported to the authorities as well as those that are not. A unique contribution of the NIS has been the use of a common definitional framework for classifying children according to types of maltreatment as well as the severity of maltreatment.

The NIS serves as the nation's needs assessment on child abuse and neglect. It offers a unique perspective on the scope of the problem beyond the children that CPS agencies investigate. While the NIS includes children who were investigated by CPS agencies, it also obtains data on other children who were not reported to CPS or who were screened out by CPS without investigation. These additional children were recognized as maltreated by community professionals. Thus, the NIS estimates include both abused and neglected children who are in the official CPS statistics and those who are not.

More information is available at https://www.acf.hhs.gov/opre/research/project/national-incidence-study-of- child-abuse-and-neglect-nis-4-2004-2009.

Periodicity

The NIS has been conducted four times – first in 1979/1980, second in 1986, third in 1993, and fourth in 2005/2006. Each collection was mandated by congress. The mandates were Public Law 93-247, Public Law 98-457, the Child Abuse Prevention, Adoption, and Family Services Act of 1988, and the Child Abuse, Domestic Violence, Adoption and Family Services Act of 1992, respectively.

ACF is committed to conducting a program of research designed to meet CAPTA requirements related to research on the national incidence of child abuse and neglect. ACF recently completed a design options study that explored innovative methods and more cost-effective strategies for addressing the range of research questions outlined in the CAPTA legislation. Two feasibility projects emerging from this work are currently being carried out:

1) Child Maltreatment Incidence Data Linkages, available at https://www.acf.hhs.gov/opre/resource/child-maltreatment-incidence-data-linkages-in-search-of-innovative-projects; and

2) Definitions and Policies Related to the Incidence of Child Abuse and Neglect and Related Risk, available at https://www.acf.hhs.gov/opre/resource/definitions-and-policies-related-to-the-incidence-of-child-abuse-and-neglect-and-related-risk-2017-2022-overview.

Data Content and Structure

The NIS gathers information from multiple sources to estimate the number of children who are abused or neglected, providing information about the nature and severity of the maltreatment; the characteristics of the children, perpetrators, and families; and the extent of changes in the incidence or distribution of child maltreatment since the time of the last NIS. Data elements include type of child maltreatment, severity of maltreatment, demographic characteristics of maltreated children, their families, and the perpetrator's age, race, sex, employment status, education, disability status, and school enrollment. The NIS also collects drug use, mental health, and level of alcohol use of those involved.

Major Publications

4th Report to Congress:
https://www.acf.hhs.gov/sites/default/files/opre/nis4_report_exec_summ_pdf_jan2010.pdf

3rd Report to Congress available for order here:
https://www.childwelfare.gov/cart/?CWIGFunctionsaction=shoppingcart:main&CWIGFunctionspk=2

2nd Report to Congress:
https://library.childwelfare.gov/cwig/ws/library/docs/gateway/Blob/28192.pdf?r=1&rpp=-10&upp=0&w=+NATIVE%28%27IPDET+PH+IS+%27%27nis-2%27%27%27%29&m=2&order=+NATIVE%28%27year%2Fdescend%27%2910&r=1&m=16

Available Datasets

Public-use aggregate:

NIS-4 data and summary tables are available to researchers through NDACAN at http://www.ndacan.cornell.edu/datasets/dataset-details.cfm?ID=147.

The Codebook is available at https://www.ndacan.cornell.edu/datasets/datasets-list-nis.cfm.

Public-use micro:

NIS-4 data and summary tables are available to researchers through NDACAN at http://www.ndacan.cornell.edu/datasets/dataset-details.cfm?ID=147.

Interactive online analyses of NIS-4 data can be produced at https://www.nis4.org/access_nis4.html.

The Codebook is available at https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset147NIS4HTMLCodebook.html.

Restricted-use micro:

Restricted use datasets are available through NDACAN at https://www.ndacan.cornell.edu/datasets/datasets-list.cfm.

In-house micro:

None available.

Statutory and Regulatory Restrictions

None.

Feasibility to Link to Other Data

Personally identifiable information:

The dataset does not include PII that would be sufficient for data linking. The data use a child's unique identification number and data do not include SSN, date of birth, address, etc. Only the PII of sentinels was gathered, and that was eventually destroyed by the contractor.

Geographic and institutional identifiers:

State and child's county of residence.

Data it has been linked to in the past:

In the NIS report to Congress, NIS data were linked to NCANDS data to look at the overlap and non-overlap of incidence. Additionally, “In addition to the main study, the NIS–4 included several supplementary studies, three of which are germane to interpreting the main study findings. Two of the studies surveyed CPS agencies—one on their overall policies, procedures and practices (CPS Structure and Practices Mail Survey – SPM) and the second on their screening standards, to determine how they would treat referrals concerning the uninvestigated cases that sentinels identified (CPS Screening Policies Study – SPS). The third supplementary study (Sentinel Definitions Survey – SDS) surveyed community professionals who served as sentinels in the main NIS–4 in order to discern the boundaries of situations they perceive to be abuse and neglect and learn what they say they would do in response to the different types of child maltreatment situations included in the NIS.

More information is available at https://www.ndacan.cornell.edu/datasets/pdfs_user_guides/Dataset147UsersGuideCodebook.pdf.

Federal funding for programs encompassed by data:

See reports to Congress, above.

Data Quality

Validity and reliability:

NIS data are uniform across variables.

Nonresponse and coverage:

See below.

Documentation:

https://www.ndacan.cornell.edu/datasets/datasets-list-nis.cfm

Overall quality:

Data quality of the NIS is considered high by subject matter experts.

Program Scope and Budget

Size of population served annually:

More than 1.25 million children experienced maltreatment during the NIS-4 study year.

Federal funding for programs encompassed by data:

$29.9 billion per year in federal, state, and local child welfare expenditures in SFY 2016, including $7.5 billion in Title IV-E funds. Source: Child Trends, https://www.childtrends.org/wpcontent/uploads/2018/12/CWFSReportSFY2016_ChildTrends_December2018.pdf.

National Survey of Child and Adolescent Well-Being (NSCAW)

Website

https://www.acf.hhs.gov/opre/research/project/national-survey-of-child-and-adolescent-well-being-nscaw

Funding Agency and Data Ownership

Children's Bureau, OPRE, ACF/HHS

Data Gathering Agency/Contractor

RTI International

Overview

The National Survey of Child and Adolescent Well-Being (NSCAW) is a nationally representative, longitudinal survey of children and families who have been the subjects of investigation by CPS agencies. There have been two cohorts of children enrolled in the survey, which makes available data drawn from first-hand reports from children, parents, and other caregivers, as well as reports from caseworkers, teachers, and data from administrative records. NSCAW examines child and family well-being outcomes in detail and seeks to relate those outcomes to experience with the child welfare system and to family characteristics, community environment, and other factors. In September 2015, OPRE in collaboration with the Children's Bureau awarded a contract to RTI International to carry out the third cohort of NSCAW (NSCAW III). An overarching goal of NSCAW III is "to maintain the strengths of previous work, while: 1) better positioning the study to address the changing child welfare population, and 2) increasing the project's overall utility." From https://www.acf.hhs.gov/opre/research/project/national-survey-of-child-and-adolescent-well-being-nscaw.

Periodicity

Interviews at baseline and follow-up with children, parents, and non-parent caregivers (see table on next page).

Table. NSCAW Data Collection Time Frames, Sample Size, and Waves

	NSCAW I	NSCAW II	NSCAW III (anticipated)
Data Collection Time Frame	1999-2007	2008-2013	2017-2022
Sample Size	N = 6,200	N = 5,800	N = 4,500
# of Waves	5 waves	3 waves	2 waves

Data Content and Structure

The study is collecting child- and family-level data from children in the child welfare system, their biological parents, caregivers, teachers (NSCAW I and II only), caseworkers, and administrators, as well as from administrative records.

NSCAW collects information on child outcomes of health and physical well-being, cognitive and school performance, mental health, behavior problems, and social functioning and relationships; caregiver health, functioning, and resources; and service needs and receipt for both children and caregivers.

Major Publications

OPRE published 21 research briefs and numerous reports apart of the NSCAW II Baseline series. Major NSCAW OPRE reports are available at https://www.acf.hhs.gov/opre/research/project/national-survey-of-child-and-adolescent-well-being-nscaw.

Available Datasets

Public-use aggregate:

None.

Public-use micro:

None.

Restricted-use micro:

NDACAN makes Restricted Data available to eligible members of the research community for the purpose of statistical analysis and reporting. NSCAW I and II data are available by request at http://www.ndacan.cornell.edu/datasets/request-nscaw-restricted-release.cfm. The two-step request process requires an application for restricted data, data protection plan, NSCAW restricted data license, and IRB approval.

In-house micro:

OPRE has all rights and access to the NSCAW dataset. Internal use of the dataset could be arranged through RTI but the standard process is to go through NDACAN. OPRE also has the ability to share data with contractors through data-sharing agreements. Any shared PII is constrained by consent form limitations.

State/local datasets:

None.

Statutory and Regulatory Restrictions

On its own, NSCAW does not have any statutory restrictions on data collection. However, NSCAW has been linked with Medicaid data in the past, and once done so, HIPAA applies. NSCAW I and II were mandated by Congress but NSCAW III is not.

Feasibility to Link to Other Data

Personally identifiable information:

Survey data contain PII, which in many cases includes SSN, date of birth, and location.

Geographic and institutional identifiers:

NSCAW I: The eight largest states are identified. NSCAW II: No states are identified. NSCAW III: TBD.

Data it has been linked to in the past:

Medicaid, NCANDS, and AFCARS. The team is considering also linking to income-related administrative data from the SSA and/or NDNH during NSCAW III.

Example research products using linked data:

Raghavan, R., Brown, D., Allaire, B., Garfield, L, Ross, R., and Snowden, L. (2014). Racial/ethnic differences in Medicaid expenditures on psychotropic medications among maltreated children. Child Abuse and Neglect 38 (2014) 1002-1010. https://www.ncbi.nlm.nih.gov/pubmed/24646610

All publications using NSCAW data may be found at https://www.zotero.org/groups/candl/items/tag/NSCAW.

Potential Research Areas to Inform Policy

It may be helpful to link NSCAW data to criminal records or to records that indicate biological non-custodial parents.

Data Quality

Validity and reliability:

NSCAW is considered a highly uniform dataset. Each iteration of NSCAW has been enhanced and modified using lessons learned from the previous cohort.

Nonresponse and coverage:

NSCAW is able to conduct extensive non-response analyses because we have access to a lot of child welfare information at sampling. There is a coverage issue in that states are excluded if they require that child welfare personnel make the first contact for consent.

In the past, NSCAW has over sampled for infants in the child welfare system. NSCAW III is proposing to oversample for infants *and* adolescents, particularly those in out-of-home care.

Documentation:

NSCAW General Release data and supporting metadata and documents are available to most researchers and require submission of the NDACAN Terms of Use Agreement. See https://www.ndacan.cornell.edu/datasets/request-nscaw-information.cfm.

Access to NSCAW Restricted Release data and supporting documents are limited to personnel at institutions that have an Institutional Review Board/Human Subjects Review Committee (IRB).

Please note: Employees at child welfare agencies are not presently eligible to obtain any version of the NSCAW data.

Overall quality:

Data quality of the NSCAW is considered high by subject matter experts.

Program Scope and Budget

Size of population served annually:

During FY 2017, approximately 3,501,000 children received an investigation or alternative response. Of those, 674,000 were determined as victims of maltreatment. Source: https://www.acf.hhs.gov/cb/resource/child-maltreatment-2017.

Federal funding for programs encompassed by data:

$29.9 billion per year in federal, state, and local expenditures in SFY 2016, including $7.5 billion in Title IV-E funds. Source: Child Trends, https://www.childtrends.org/wp-content/uploads/2018/12/CWFSReportSFY2016_ChildTrends_December2018.pdf.

National Survey of Children in Non-Parental Care

Website

https://www.cdc.gov/nchs/slaits/nscnc.htm

Funding Agency and Data Ownership

Funding provided by ASPE and The Annie E. Casey Foundation. Data are owned by the National Center of Health Statistics at the Centers for Disease Control and Prevention (CDC).

Data Gathering Agency/Contractor

CDC's National Center for Health Statistics (NCHS), State and Local Area Integrated Telephone Survey program, under contract with NORC, University of Chicago

Overview

This survey provides nationally representative estimates on the characteristics, living arrangements, and service accessibility of noninstitutionalized children who were living apart from their parents (in foster care, grandparent care, or other nonparental care) and who were aged 0 to 16 years in 2011-2012. Data on the well-being of the children and of their caregivers are also available. The children's nonparental care status was identified in a previous SLAITS survey, the 2011-2012 National Survey of Children's Health.

Major goals of the data collection are to:

1) Describe the characteristics of the population of children living apart from their parents;
2) Assess the health and well-being of children in nonparental care and of their caregivers; and
3) Assess access to and utilization of various supports and services by these children and their caregivers. The survey has a sample size of 1,298.

More details can be found at https://www.cdc.gov/nchs/data/slaits/nscnc_faq_and_guidelines_for_data_users.pdf.

Periodicity

One time (2013).

Data Content and Structure

- Whether the child still lived in the same household as at the time of the NSCH survey.
- If the child no longer lived in the NSCH household, circumstances related to how and when the child started and stopped living in that household (the "No Longer Nonparental Caregiver" – or NLNPCG – Module).
- Current housing and living arrangements.
- Child's contact with siblings.
- Characteristics of the child's parents (includes drug/alcohol problems, mental health problems).
- Child's and child's caregiver's interactions with child's parents.
- Financial support, including child support from the parents, foster care subsidies, other governmental assistance, and non-monetary (in-kind) support.
- Caregiver preparation, long-term planning, and legal custody.
- Child well-being, child care, and school.
- Caregiver health, well-being, and social supports.
- Child's health insurance and caregiver's health insurance.
- Household income and program participation.
- Demographic and socioeconomic characteristics of the child's caregiver and caregiver's spouse.

Source:
https://www.cdc.gov/nchs/data/slaits/nscnc_faq_and_guidelines_for_data_users.pdf.

Major Publications

ASPE has produced a number of descriptive briefs and reports, including:

- https://aspe.hhs.gov/pdf-report/children-living-apart-their-parents-highlights-national-survey-children-nonparental-care
- https://aspe.hhs.gov/pdf-report/complexities-kinship-care-key-findings-2013-national-survey-children-nonparental-care
- https://aspe.hhs.gov/pdf-report/children-living-apart-their-parents-highlights-national-survey-children-nonparental-care

Available Datasets

Public-use aggregate:

None.

Public-use micro:

Two files are included in the public release: the NLNPCG Screener File and the NSCNC Interview File. See ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/slaits/nscnc/2_Datasets.

Restricted-use micro:

It is possible to conduct analysis of linked NSCH/NSCNC data. Each child in the NSCNC also has an NSCH record. However, because the NSCH was designed as a state-level file and includes state identifiers on the public data file, while NSCNC was designed as a national file that is not representative at the state level, it is not possible to link the public NSCH and public NSCNC files. Analysts interested in using the linked data should apply to the NCHS Research Data Center for access to nonpublic data. Please see the procedures at http://www.cdc.gov/rdc/. The NSCH adds demographic, health, and well-being variables that were not repeated in the NSCNC and therefore do not appear in the NSCNC public use file.

In-house micro:

Federal staff interested in analyses that require access to restricted data may contact Matthew Bramlett for additional information at Mbramlett@cdc.gov.

Statutory and Regulatory Restrictions

Data may be used for research purposes only. The collection of the data is authorized under Section 306 of the Public Health Service Act (Title 42, US Code, Section 242k). Confidentiality is assured under Section 308d of this Act. Only NCHS employees or agents may have access to PII.

Source: Survey instrument, page 10, https://www.cdc.gov/nchs/data/slaits/nscncquestionnaire.pdf.

Feasibility to Link to Other Data

Personally identifiable information:

Name and telephone number are collected and available to CDC/NCHS staff for limited allowable research purposes (e.g., for the administration of follow-back surveys).

Geographic and institutional identifiers:

State identifiers are not available through the public dataset, and are only available on a restricted basis through the NCHS Research Data Center. See http://www.cdc.gov/rdc/.

Data it has been linked to in the past:

It is possible to conduct analysis of linked NSCH/NSCNC data. [NSCH = National Survey of Children's Health]. Because the NSCH was designed as a state-level file and includes state identifier on the public data file, while NSCNC was designed as a national file that is not representative at the state level, it is not possible to link the public NSCH and public NSCNC files together. Analysts interested in using the linked data should apply to the NCHS Research Data Center for access to nonpublic data. Please see the procedures at http://www.cdc.gov/rdc/.

Example research products using linked data:

Children Living Apart from their Parents: Highlights from the National Survey of Children in Nonparental Care. https://aspe.hhs.gov/system/files/pdf/203352/NSCNC.pdf.

Bramlett, M and Radel, L. (2016). Factors Associated With Adoption and Adoption Intentions of Nonparental Caregivers. Adoption Quarterly, Vol 20 (1), 5-24.

Bramlett, MD, Radel, LF and Chow, K. (2018). Health and Well-Being of Children in Kinship Care: Findings from the National Survey of Children in Nonparental Care. *Child Welfare 95(3):40-60*.

Potential Research Areas to Inform Policy

Questions that could be explored:

- In what ways do differences in families' level of child welfare involvement (e.g., whether the child is in the custody of the child welfare agency or whether there has ever been an open child welfare case) relate to the well-being of children who live with neither parent?
- Are particular forms of continued parental involvement in the lives of children with whom they do not live related to children's characteristics (e.g., age, length of time since they lived with the parent) and/or their children's health and well-being?

Data Quality

Validity and reliability:

Most data elements appear to be reasonably reliable. One exception is data on SSI receipt. Those numbers are very high and it is believed some grandparent respondents may have misunderstood the question and answered including their own social security receipt and not just children's payments. Some data about custody and foster care also seem inconsistent. It may be that some respondent caregivers were not entirely clear about their custody status, or confused legal custody with physical custody.

Source: ASPE staff.

Nonresponse and coverage:

The response rate is the number of completed interviews as a proportion of the number of eligible units in the sample. For NSCNC, this rate is the product of the recontact rate (70.4 percent) and the interview completion rate among recontacted households (74.5 percent), or 52.4 percent. However, the NSCH response rate was 23.0 percent. Thus, accounting for nonresponse to the NSCH, the overall NSCNC response rate is the product of the NSCNC response rate and the NSCH response rate, or 12.1 percent. It should be noted that the NSCH response rate of 23 percent does not mean that three-quarters of eligible households refused to participate in the NSCH. The NSCH response rate is low in part because of cell phone numbers that ring with no answer and for whom eligibility cannot be determined. The NSCH cooperation rate among eligible households, or interview completion rate, was 51.4 percent.

Documentation:

A Guidebook, Instrument, Data Dictionary with frequencies, SAS Programs, and the datasets themselves are all available at https://www.cdc.gov/nchs/slaits/nscnc.htm.

Overall quality:

Data quality of the NSCH is considered high by subject matter experts.

Program Scope and Budget

State/local datasets:

Datasets with state identifiers are available on a restricted basis through the NCHS Research Data Center.

Size of population served annually:

Around 3 million children live in households with no parents (3.9 percent of the child population). Source: https://aspe.hhs.gov/system/files/pdf/76911/rpt_nonparentalcare.pdf.

Federal funding for programs encompassed by data:

Not applicable.

National Survey of Early Care and Education (NSECE)

Website

https://www.acf.hhs.gov/opre/research/project/national-survey-of-early-care-and-education-nsece-2010-2014

Funding Agency and Data Ownership

ACF, HHS

Data Gathering Agency/Contractor

NORC at the University of Chicago

Overview

The National Survey of Early Care and Education (NSECE) documents the nation's current utilization and availability of child care and early education (including school-age care) in order to deepen the understanding of the extent to which families' needs and preferences coordinate well with providers' offerings and constraints. The experiences of low-income families with young children are of special interest, as they are the focus of a significant component of early care and education/school-age (ECE/SA) public policy. The NSECE collects data through a set of integrated, nationally representative surveys including interviews in all 50 states and Washington, D.C.

The NSECE 2012 includes five survey components and four related questionnaires:

1) A Household Survey conducted with a parent or guardian of a child or children under age 13. Eligible respondents were identified through the Household Screener. In 2012, the NSECE data include approximately 12,000 interviews with adults in households with children under age 13.

2) A Home-based Provider Survey conducted with two types of respondents. The first type includes formal Home-based Providers who were identified through state-level administrative lists (of ECE/SA providers) as providing regulated or registered home-based care, with an estimated total of 4,000 interviews. The second type includes unlisted Home-based Providers identified through the Household Screener as caring for children under age 13 who are not their own in a home-based setting (and who do not appear on a state-level administrative list), with an estimated total of 2,000 interviews.

3) The Center-based Provider Survey conducted with directors of ECE/SA programs identified from state-level administrative lists such as state licensing

lists, Head Start program records, or pre-K records. Eligible respondents were identified through the Center-based Provider Screener. The estimated total of Center-based Provider interviews is 8,200.

4) The Workforce Survey was conducted with classroom-assigned staff members of Center-based providers who completed the Center-based Provider interview. After each Center-based Provider interview was completed, one staff member from that organization was sampled and administered the workforce interview. Approximately 5,600 Center-based Workforce members were interviewed. In addition, the Home-based Provider questionnaire collected workforce information from those providing child care in home-based settings. [https://www.acf.hhs.gov/opre/research/project/national-survey-of-early-care-and-education-nsece-2010-2014]

The five components of the NSECE will be fielded again in 2019.

Periodicity

2012, 2019.

Data Content and Structure

Household Survey:
This survey documents the nation's demand for child care and early education services. Key questionnaire topics include details on usage of non-parental care, expenditures on non-parental care, preferences and search activities for child care, and the balance of parental employment with child care needs and availability. Data from multiple children, details of parental searches for care, parent and child schedules during the week prior to the interview, and innovative approaches for determining likely participation in government programs (such as CCDF, Head Start, or public pre-K) are all innovations in the household survey.

Home-Based Provider Survey:
Key questionnaire topics in the home-based provider questionnaire include enrollment and the characteristics of the children served, rates charged for care, characteristics of providers and care provided, household characteristics, qualifications for and attitudes about child care, activities conducted with children, and participation in government programs.

Center-Based Provider Survey:
Topics covered by this instrument include enrollment and characteristics of children served, staffing, prices charged, schedules of service, participation in government programs, and staff compensation and professional development policies and activities. The questionnaire also includes the selection of a representative classroom about which

more detailed staffing, compensation, and curriculum information are collected. Although no observational data are collected, the questionnaire includes a variety of measures at both the program and individual staff levels that have been found in the literature to predict observed quality of care.

Workforce (Classroom-Staff) Survey:
Topics include information about the work setting (e.g., activities in the classroom, interactions with parents and other staff, availability of professional development and other supports), roles and responsibilities (lead teacher, teacher, assistant teacher, aide), compensation (wages and benefits), and perceived leadership and morale, as well as personal information about qualifications, attitudes about child care, and stress, depression, and demographic information.

(From ICPSR: https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35519#funding.)

Major Publications

Several reports are available on the Homepage. Additional OPRE-funded projects produce reports using the NSECE data from time to time, including https://www.acf.hhs.gov/opre/research/project/child-care-and-early-education-policy-and-research-and-technical; https://www.acf.hhs.gov/opre/research/project/center-for-research-on-hispanic-children-families; and https://www.acf.hhs.gov/opre/secondary-analysis-of-data-on-child-care-and-early-education-2015-2017.

Available Datasets

Public-use aggregate:

Quick Tabulation files (workforce survey, center-based provider survey, unlisted home-based provider survey, listed home-based provider survey, household survey, and household child-level data), including codebooks, are available through ICPSR. See https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35519.

OPRE also released a book of tables created by NORC summarizing usage and cost of non-parental care, available at https://www.acf.hhs.gov/opre/resource/early-care-education-usage-households-out-of-pocket-costs-tabulations-nsece.

Public-use micro:

Public Use Datasets (workforce survey, center-based provider survey, home-based provider survey, household survey, household calendar data), including codebooks, are available through ICPSR at https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35519.

Restricted-use micro:

Three levels of restricted-use data exist for the NSECE:

Level 1 Restricted-use data are available by request accompanied by IRB approval through ICPSR at https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35519.

Level 2 Restricted-use data include geographic identifiers and are available by requesting an application from NORC at nsece@norc.org.

Level 3 Restricted-use data include identifying information that is available to authorized staff under stringent conditions.

In-house micro:

None.

Statutory and Regulatory Restrictions

Section 6580 (a)(5) of the CCDBG Act of 2014 authorizes the Secretary of HHS to use funds for research and demonstration activities as well as periodic external, independent evaluations of the impact of the program described by this subchapter on increasing access to child care services and improving the safety and quality of child care services using scientifically valid research methodologies, and to disseminate the key findings of those evaluations widely and on a timely manner.

Feasibility to Link to Other Data

Personally identifiable information:

Respondents to the Household Survey were asked to consent to link children’s data to administrative records. If the respondent consented, full names, dates of birth, and an address were collected for all children in the household. These data are maintained as Level 3 Restricted-use files by NORC.

Geographic and institutional identifiers:

Level 2 Restricted- use files maintained by NORC include geographic identifiers at the county (or county-cluster) and state levels. The name and address of non-parental care providers as well as the address of the place of school/work/training for household caregivers are included in Level 3 Restricted-use files maintained by NORC.

Data it has been linked to in the past:

NSECE household data were linked with Illinois CCDF administrative records. NSECE provider data have been linked with state policies, notably using the CCDF Policies Database.

Example research products using linked data:

Methodology report on linking with IL state data: https://www.acf.hhs.gov/opre/resource/examining-child-care-subsidy-receipt-an-analysis-of-matched-nsece-and-illinois-administrative-data.

Potential Research Areas to Inform Policy

Linking the NSECE household data with OCC's 801 data on children receiving CCDF subsidies would enable in-depth analysis of characteristics of families who are eligible but do not receive CCDF subsidies.

Data Quality

Validity and reliability:

Overall the NSECE data quality is good.

Nonresponse and coverage:

Weighted response rates (screening interview response rates) were: 62.2 percent (Household Survey); 80.7 percent (Listed Home-Based Provider Survey); 67.5 percent (Unlisted Home-Based Provider Survey); 73.7 percent (Center-Based Provider Survey); and 71.2 percent (Workforce Provider Survey). Item non-response was not a particular concern.

Documentation:

Numerous resources, including a methodology report, are available through OPRE and ICPSR. See https://www.acf.hhs.gov/opre/research/project/national-survey-of-early-care-and-education-nsece-2010-2014 and https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35519.

Overall quality:

Data quality of the NSECE is considered high by subject matter experts.

Program Scope and Budget

State/local datasets:

Level 2 Restricted-use data can be linked to state or county (or county-cluster) identifiers by requesting an application from NORC at nsece@norc.org.

Size of population served annually:

In 2012, over 15 million children age 5 or younger participated in regular non-parental care. This care was provided by 129,000 center-based programs, 118,000 listed home-based programs, and by over 3,000,000 other non-parental caregivers (unlisted home-based providers).

Federal funding for programs encompassed by data:

In FY 2016 the federal government expended approximately $5.68 billion for the CCDF program, approximately $9.3 billion for Head Start, and $250 million for Preschool Development Grants.