

September 2011

Strategies for Leveraging SNACC Data for Policy and Evaluation: Sustainability of Data Linkage Projects and Coalitions of Interest

-Third in the Series, "The Next Generation of Data Linkage Projects"

Michael J. O'Grady, Ayesha Mahmud

NORC • University of Chicago • Chicago, IL

INTRODUCTION

The Importance of Data Linkages

The demand for data for the purposes of policy analysis has increased dramatically, with health services researchers and policymakers alike poised to analyze and evaluate a range of challenges to the health care system. However, health services researchers and policymakers often face the challenge of having incomplete data. Health-related data, whether from surveys, claims, or administrative records, are often created and held by different public and private entities. To address this disconnect, individual data sets can be linked to one another, providing a more comprehensive overarching data set while avoiding the cost of duplicate data collection.

The process of linking data sets can take a number of different forms, but the common characteristic of the process is the application of statistical methods in order to identify and connect the same—or demographically similar—individuals within each of the data sets. For example, a researcher might match survey respondents with their actual claims and eligibility files in order to study the correlation between income level and service utilization.

numbers reported in state and national administrative data. With the financial support of the Robert Wood Johnson Foundation, six organizations (see text box below) joined forces to conduct six data linkage projects to determine which data source provided the most robust estimates of the Medicaid population. These projects are listed in Table 1. In addition to the linkages established under SNACC, several other useful data linkage projects have been conducted. These non-SNACC projects are listed in Table 2.

SNACC Organizations

- State Health Access Data Assistance Center (SHADAC)
- National Center for Health Statistics (NCHS)
- Agency for Healthcare Research & Quality (AHRQ)
- Assistant Secretary for Planning & Evaluation (ASPE)
- Centers for Medicare & Medicaid Services (CMS)
- Census Bureau

SNACC has always had a focus on informing policy: The larger aim of identifying the most robust Medicaid estimates was to provide policy makers with more accurate approximations of the Medicaid and uninsured populations in order to facilitate the creation of effective policy. The project then evolved into a [powerful] set of analytic files uniquely positioned to inform policy development, implementation, and evaluation.

BACKGROUND

SNACC

SNACC—alternately referred to as the “Medicaid Undercount Project”—began as a collaborative effort to explain why discrepancies exist between survey estimates of Medicaid enrollment and the enrollment

DATA LINKAGES

Sustainability of Linkage Projects and Coalitions of Interest

Triggered by the pressures of implementing the Affordable Care Act (ACA), both the federal and state governments need robust data in order to effectively meet the challenges of designing new programs, implementing those designs, and evaluating the outcomes. This brief is the third in a three-part series, “The Next Generation of Data Linkage Projects.”

The first brief explored priority areas for data linkages, looking at health benefit exchanges, the intersection of

Medicare and Medicaid, and Medicare payment reform. The second brief analyzed three challenges to creating linked data sets: (1) methodological challenges; (2) privacy concerns; and (3) barriers to data access. The present brief discusses potential coalitions of interest inside and outside the federal government that can sustain data linkage projects over time by providing financial support as well as political and policy support. The ideas outlined here were identified during interviews with key current and former policy makers and analytic staff both inside and outside of government, along with a thorough review of relevant literature, with a primary focus on informing implementation, design, and evaluation.

Table 1:

The SNACC Project: Data Linkages to Date

The national-level CPS database
The Medicaid Statistical Information System (MSIS) and the Current Population Survey (CPS)
The state-frame, household, and person MSIS data to the CPS
The MSIS and the National Health Interview Survey (NHIS)
The MSIS and the CPS annual Social and Economic Supplement (ASEC), 2003-2004
The MSIS and the Medical Expenditure Panel Survey (MEPS)

Table 2:

Other Data Linkage Initiatives

Health insurance data from the 2001 State and Local Area Integrated Telephone Survey's National Survey of Children with Special Health Care Needs (NS-CSHCN) linked to immunization status data from the 2000-2002 National Immunization Survey (NIS)
Cancer registry data from the Surveillance, Epidemiology, and End Results (SEER) linked to Medicare managed care enrollee survey data from the Medicare Health Outcomes Survey (MHOS)
Tract-level poverty data linked to vital records for infants born to American Indian women between 1990 and 1999
Administrative data from the Medicaid/State Children's Health Insurance Program linked to birth and death records
Air pollution data from the National Health Interview Survey (NHIS) linked to data from the National Hospital Discharge Survey (NHDS)
Survey data from the National Center for Health Statistics (NCHS) linked to death certificate records from the National Death Index
Survey data from NCHS linked to claims data from CMS
Survey data from NCHS linked to benefit records from the Old Age, Survivors and Disability Insurance (OASDI) and Supplemental Security Income (SSI)

The Current Situation

At present, the federal statistical agencies—in particular the U.S. Census Bureau, the National Center for Health Statistics (NCHS), and the Centers for Medicare and Medicaid Services (CMS)—are conducting most of the data linkage work involving national datasets. It is the statistical staffs at these agencies, rather than federal policymakers in either the executive or legislative branches, that are driving most of this work. The leadership at these agencies and at the federal Office of Management and Budget (OMB) are supportive of linkage efforts and have provided what resources they are able to provide; however, without a clear champion among the political leadership in either branch of government, there has been no aggressive pursuit of funding for data linkages as an important policy issue. This political championship is lacking because, while the statistical staffs can see the powerful potential of linked datasets, the policymakers are still unconvinced because they have yet to see a clear demonstration of the potential promised by statistical staffs.

On the positive side, the lack of political interest in data linkages has afforded the federal statistical staffs the opportunity to work through the many methodological challenges in building linked data sets (e.g., revised sampling weights and standard errors) without the political pressure to produce results for ongoing policy debates. However, without the backing of policymakers, the continued viability of the data linkage process is questionable given the current budget climate.

It appears that the ACA, with its implementation and design challenges, might provide the impetus to carry data linkage efforts to the next level. These challenges are significant and would seem to necessitate powerful analytic tools like linked datasets. This need is particularly clear for implementation tasks such as establishing state-level health insurance exchanges and controlling costs for the beneficiaries eligible for both Medicare and Medicaid (i.e., the “dual eligibles”). It remains to be seen, though, if policymakers responsible for ACA implementation will see the potential of linked data sets. However, there is evidence of some increased interest in data linkages, and it is likely that the analytic support agencies in the legislative branch (e.g., the Congressional Budget Office, the Congressional

Research Service, the Government Accountability Office, and the Medicare Payment Advisory Commission) will begin to use linked data sets more and more for budget estimates and reports to Congress. This use would certainly help raise awareness of the importance of data linkages as analytical tools in the policymaking process. However, an effort to proactively raise awareness about the importance of data linkages will likely be required: Policymakers are often ignorant of the extent to which data activities in general contribute to their essential policy goals despite the significant contribution of, for example, the Medical Expenditure Panel Survey (MEPS) to CBO cost estimates, or the role of the National Health and Nutrition Examination Survey (NHANES) in national projections from the Centers for Disease Control and Prevention (CDC) of clinical trial results. As a result of this ignorance, such surveys are chronically underfunded, and seldom involve anything as statistically nuanced as data linking.

Possibilities for the Future

With the exception of the Robert Wood Johnson Foundation’s participation in the SNACC project, it appears that no other private sector entity is actively working on data linkage projects. There are no state efforts in place, either, except for one project undertaken in Maine. Other than these two exceptions (RWJF and Maine), the effort to create data linkages is occurring almost entirely at the federal level. The reason for this focus is not only financial; many of the more sophisticated adjustments that occur in the process of linking data require access to federal microdata on individuals, to which the Census Bureau alone has legal access.

Future funding of linkage projects could take the form of federal interagency collaborative efforts to fund specific projects. This type of funding structure would parallel that of most of the major federal surveys: Either one agency would take the lead in both content and funding, as CMS does for the Medicare Current Beneficiary Survey (MCBS); or a consortium of agencies would contribute, as with the Centers for Disease Control and Prevention (CDC) State and Local Area Integrated Telephone Survey (SLAITS), for which different HHS agencies fund specific questions. Given the HHS decentralized data policy apparatus, the particular funding arrangement used would likely depend upon which agency has questions they need

answered and which agency has the money to fund the work. Unfortunately, either data linkage funding arrangement would probably suffer from the same downsides seen with the parallel survey funding arrangements: a lack of overarching data policy at the departmental level; competition, rather than cooperation, between statistical agencies; and poor funding of methodological research.

While direct funding for any future data linkage efforts will be publicly-sourced, private sector groups could have a significant impact on these efforts in a number of indirect ways. For example, private funding for methodological research would provide a critical service. A more technically sophisticated linkage methodology would help avoid estimation errors, which, if relevant to a question of intense political debate, could be a significant setback to the adoption of linked data sets. Funding for indirect research such as this could be distributed on a grant-by-grant basis to outside researchers, as is the traditional route. Alternately, given privacy constraints on who can have access to data, funding of a fellowship for researchers to work directly at the Census Bureau and other federal statistical agencies might be a more effective means of supporting these efforts.

CONCLUSION

The future of data linkage projects is currently at a crossroads. The first generation of these projects is complete, and much has been learned to date about the technical and methodological nuances of linking different data sets. The Affordable Care Act presents a unique policy challenge, and data linkage can provide a powerful new tool for ACA implementation. Whether policymakers will realize the potential of data linkages and gravitate toward this tool is still an open question, but support from the policy realm will be critical in order to ensure that linkage efforts move forward to the next stage. Backers of data linkage will have to be focused in their support and stress the value of these programs and their significant return on investment.

This brief is a companion to the brief titled, “The Next Generation of Data Linkage Projects: Priority Areas for SNACC under the ACA,” which is available at http://www.shadac.org/files/shadac/publications/SNACCDDataLinkageBrief_1of3.pdf, and to the brief, “Strategies for Leveraging SNACC Data for Policy and Evaluation: Barriers and challenges to Linked Data Sets,” which is available at http://www.shadac.org/files/shadac/publications/SNACCDDataLinkageBrief_2of3.pdf.

ACKNOWLEDGMENTS

The authors would like to thank the following people for their time and insights: Richard Kronick, PhD, Deputy Assistant Secretary for Health Policy (ASPE); Linda Bilheimer, PhD, Associate Director Analysis and Epidemiology (NCHS); Chris Peterson, MPP, Director of Eligibility, Enrollment and Benefits (MACPAC); April Grady, MPAff, Principal Analyst (MACPAC); John Poisal, Deputy Director National Health Statistics Group (CMS); David Baugh (CMS); Stephen Heffler, Director National Health Statistics Group (CMS); Christine Cox, Director Research and Information Dissemination (CMS); Kimberly Lochner (CMS); William Scanlon, PhD, Health Policy Consultant; Amy O’Hara, PhD, Data Integration Division, US Census Bureau; Sharon Arnold (OCIIO); Cara Lesser (OCIIO); Helen Levy, PhD (Council of Economic Advisors)

ABOUT SHARE

The State Health Access Reform Evaluation (SHARE) is a Robert Wood Johnson Foundation (RWJF) program that supports rigorous research on health reform issues, specifically as they relate to the state implementation of the Affordable Care Act (ACA). The program operates out of the State Health Access Data Assistance Center (SHADAC), an RWJF-funded research center in the Division of Health Policy and Management, School of Public Health, University of Minnesota. Information is available at www.statereformevaluation.org.

State Health Access Data Assistance Center
2221 University Avenue, Suite 345
Minneapolis, MN 55414
Phone (612) 624-4802