

September 2011

The Next Generation of Data Linkage Projects: Priority Areas for SNACC under the ACA

-First 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.

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 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.

DATA LINKAGES UNDER THE ACA

With the Affordable Care Act (ACA) in place, 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 new health policy landscape calls for an examination of priority areas for

the next generation of SNACC linkage projects. Interviews with key current and former policy makers and analytic staff both inside and outside of government, a thorough review of relevant literature, and an analysis of the demands of health care reform implementation, point to potential linkages that would be both relevant and useful in a post-reform environment.

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)

Where Linked Data Sets Would Be Beneficial

The Design and Implementation of the ACA's Health Benefit Exchanges

Exchange design and implementation could benefit from a comprehensive data system that has both public and private data on coverage, eligibility, benefits, and costs.

The ACA envisions a state-level Federal Employee Health Benefits Plan (FEHBP) – style multi-plan system. As with FEHBP, either the state or federal government will manage the exchange, carrying out administrative tasks including: the determination of which plans are eligible for inclusion in the exchange; the determination of which applicants are eligible to purchase coverage through the exchange; and the determination and distribution of premium subsidies for eligible applicants. The private sector insurers will provide the actual insurance coverage, negotiate prices with providers, etc. However, in contrast to the FEHBP, employers will be able to decide whether to participate in the exchange or not. Additionally, there will be an “essential benefits” package that all plans—whether operating inside or outside the exchange—will have to provide as a minimum actuarial value.

An FEHBP-style exchange design like that described here presents a number of challenges, and policymakers will need proper data systems in order to effectively meet these challenges. Consider, for example, the question of adverse selection. Systems like FEHBP and Medicare Advantage have decades of experience with adverse selection and the tools used to counteract its effect (e.g., risk-adjusted premiums and constrained actuarial value). The ACA exchanges will pose a new problem: the potential for employer-level adverse selection. Namely, employers whose workforces are more expensive to insure than the exchange average will have a strong incentive to move their people into the exchange to reduce costs. On the other hand, employers whose workforces are less expensive than the exchange average will have a strong incentive to keep their people *out* of the exchange to reduce costs. In essence, the exchange could evolve into a state high-risk pool within a few years time unless exchange administrators put into place effective measures to counteract adverse selection.

To guard against the potential for adverse selection, exchange designers and administrators will need robust data on a number of variables. For example, in order to construct an accurate risk adjustment methodology and provide guidance on appropriate benefit variation, exchange planners will need data on the expected expenses of the populations to be insured and on the benefits to be offered. Planners will also require information about the cost-sharing requirements, network adequacy, and other characteristics of plans both inside and outside the exchange. Some of this data is already collected; however, it is currently gathered by multiple different organizations and in different formats. A combination of current survey and administrative data from the Department of Health and Human Services, the Department of Labor, and the Department of the Treasury could help answer many of the questions that exchange planners will need to answer.

The Intersection of Medicare and Medicaid

The complex intersection of Medicare and Medicaid—where we find those who are simultaneously eligible for both programs, i.e., the “dual eligibles” –could be better understood by linking socioeconomic data with administrative data; by linking Medicaid long- term care data with home health care data; and by linking provider information to claims data.

The ACA establishes an enhanced federal role in the care of dual eligibles that focuses on quality, access, and cost. This population is both extremely vulnerable and extremely expensive, so the need to effectively coordinate their care is a high priority. In order to redesign a system of care that will meet the needs of this population in a manner that is as careful and cost-effective as possible, analysts and policymakers need to gain a comprehensive picture of the health care needs of this group, their patterns of care utilization, and the mix of providers and settings used. To facilitate this understanding, the information resources from both Medicare and Medicaid have to be linked in a robust way. Data sources coming from surveys, claims, and administrative records for both programs are essential to the establishment of a comprehensive picture of the care, quality, and costs associated with an expensive subpopulation with complex health care needs.

Medicare Payment Reform

Medicare payment reform could benefit from linkages between patterns of access and utilization, because these linkages could facilitate the development of new models of care and financing. Researchers could also link readmission patterns to dual-eligibility status in order to explore the possibility of reducing costs for this group by bundling readmissions with other services. Linkages between administrative and claims data for Medicare Parts A and B and for Medicare Advantage plans, would facilitate the creation of a baseline to determine where specific changes are needed and to evaluate the success of efforts to make improvements in these areas.

However, the knowledge and use of these powerful new data sets among policymakers is still in its infancy. Moreover, current health policy challenges will require powerful new tools along the lines of SNACC's linked data sets, and it appears that there is still a significant gap between the available data and the data needs of policy analysts and policymakers in this area.

The future of data linkage efforts will ultimately depend on the development of a strong working relationship between the technical staff who produce the data and the policy staff who are the logical end-users of the data. While successful at meeting its original goal of informing policy, the first generation of SNACC data sets has become somewhat dated and therefore less germane to the policy challenges ahead. In order to meet the current needs of federal and state policymakers, the SNACC project must move on to its second phase. This brief identifies a few of the most pressing applications of linked data sets associated with the design, implementation, and reform of health programs in the United States.

CONCLUSION

In the realm of data linkage efforts, both positive and negative developments can be seen. The technical work on building linked data sets is moving forward, primarily in the public sector.

ACKNOWLEDGMENTS

The authors would like to thank the following people for taking the time to be interviewed for this study: Richard Kronick, PhD, ASPE; Linda Bilheimer, PhD, NCHS/CDC; Chris Peterson and April Grady MACPAC; William Scanlon, PhD, National Committee of Vital and Health Statistics (NCVHS); Christine Cox, Jon Poisal, David Baugh, Steven Heffler and Kimberly Lochner, CMS; Amy and Bret O'Hara, Census; and Sharon Arnold and Cara Lesser, CCIIO/CMS.

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