



HHS Public Access

Author manuscript

Annu Rev Public Health. Author manuscript; available in PMC 2019 April 01.

Published in final edited form as:

Annu Rev Public Health. 2018 April 01; 39: 437–452. doi:10.1146/annurev-publhealth-040617-013544.

Data Resources for Conducting Health Services and Policy Research

Lynn A. Blewett, Ph.D.,

Professor of Health Policy and Director, State Health Access Data Assistance Center (SHADAC), Division of Health Policy and Management, School of Public Health, University of Minnesota, 2221 University Avenue, Suite 345, Minneapolis, Minnesota, 55414. 612-624-4802; blewe001@umn.edu

Kathleen Thiede Call, Ph.D.,

Professor, Division of Health Policy and Management School of Public Health, University of Minnesota, 420 Delaware Ave, S.E., MMC #729, Minneapolis, Minnesota, 55455. 612-625-6151; callx001@umn.edu

Joanna Turner, M.S., and

Senior Research Associate, State Health Access Data Assistance Center (SHADAC), Division of Health Policy and Management, School of Public Health, University of Minnesota, 2221 University Avenue, Suite 345, Minneapolis, Minnesota, 55414. 612-624-4802; turn0053@umn.edu

Robert Hest, M.P.P.

Research Associate, State Health Access Data Assistance Center (SHADAC), Division of Health Policy and Management, School of Public Health, University of Minnesota, 2221 University Avenue, Suite 345, Minneapolis, Minnesota, 55414. 612-624-4802; hestx005@umn.edu

Abstract

Rich federal data resources provide essential data inputs for monitoring the health and health care of the U.S. population and are essential for conducting health services policy research. The six household surveys we document cover a broad array of health topics from health insurance coverage (American Community Survey, Current Population Survey); health conditions and behaviors (National Health Interview Survey, Behavior Risk Factor Surveillance System); health care utilization and spending (Medical Expenditure Panel Survey); and longitudinal data on public program participation (SIPP). New federal activities are linking federal survey with administrative data to reduce duplication and response burden. In the private-sector vendors are aggregating data from medical records and claims to enhance our understanding of treatment, quality and outcomes of medical care. Federal agencies must continue to innovate to meet the continuous challenges of scarce resources, pressures for more granular data and new multi-mode data collection methodologies.

Keywords

Data; Public Health Surveillance; Household Survey Data; Health Services Research

1. Introduction

Health policy is increasingly complex as major efforts to reform the U.S. health care system are met with partisan divide over core values about the role of government involvement in health insurance coverage. The policies designed and promoted on both sides of the aisle are dependent on what we know about current levels of private and public health insurance coverage, the health care needs of the population, and the role of government in financing health care. Our national data resources serve as inputs for policy analysis and projection models that inform this intense policy debate.

In this paper, we provide an overview of key federal household surveys that are used in health services policy research. We present information on elements of each resource, describing their strengths and limitations, with the anticipation that this synthesis will be used as a type of users guide. We include a discussion of innovations in data collection including an overview of data linkage projects and relatively new efforts to monitor cost and utilization through the aggregation of patient claims data. We conclude with observations of the future of U.S. household surveys.

Approach

We focus our review on data resources that include a broad range of topics and questions and include the general population in its sample. We exclude surveys targeted to a specific topic or target population. We include surveys that are ongoing and currently funded; follow a regular schedule for data collection; and provide data access through a public-release process. We include federal administrative data only when linked to national household surveys.

The timing of this paper is critical. There are new advances in data collection and data technology that are pushing the field of survey research to think creatively and to design surveys that are useful, efficient, and minimize response burden. The future of the federal surveys and innovations in data collection and linkages will be critical to the ability of health services research to inform health policy.

2. Key Federal General Population Health Surveys

Health services researchers are fortunate to rely on a range of consistent sources of data available to monitor the impact of social, economic, and health related policies. Table 1 provides an overview of the health related content included in each survey included in our review. The amount of health related content varies depending on the main purpose of each survey. The NHIS, BRFSS, and MEPS-HC provide the richest sources of health-related content. However, the NHIS and MEPS-HC has more limited sample sizes that constrain state-level analysis (see Table 2).

American Community Survey (ACS)

The American Community Survey (ACS) (54) is a general household survey that was designed as a replacement for the long form of the decennial Census. The ACS provides annual estimates on demographic, socioeconomic, and housing information. The ACS is a

continuous mixed-mode survey with mail and internet (added in 2013) collection with phone and in-person used for non-response follow-up. The ACS is a mandatory survey, although there are efforts to make the survey voluntary (16). This would likely have a negative impact on response rates and increase costs (87, 105).

The greatest strength of the ACS is its large sample size and ability to produce estimates for states and sub-state geographies, such as congressional districts, counties, and zip-code tabulation areas. The ACS samples about 3.5 million addresses annually (94). The key limitation for health services research is that it does not have in-depth health related content. Every question on the ACS requires federal justification (99), which keeps the ACS from responding quickly to changes in policies. Proposed question changes must be approved by the Office of Management and Budget (OMB) and the Interagency Council for Statistical Policy and go through a content test before implementation (93).

The ACS did not make any updates in its questionnaire to accommodate health reform, unlike the other federal surveys discussed in this paper. Questions on exchange participation and subsidies are currently being evaluated on the 2016 Content Test (88), but, with the current political debates involving health reform, these questions might be outdated by the time they are ready for implementation.

Current Population Survey (CPS)

The Current Population Survey (CPS) (95) provides monthly data regarding labor force participation and unemployment for the civilian non-institutionalized population. It is conducted by the U.S. Census Bureau on behalf of the Bureau of Labor Statistics. The Annual Social and Economic Supplement (ASEC) collects data by phone and in-person, on income and health insurance coverage, from February to April of each year. The CPS has additional supplements fielded throughout the year on a variety of topics (107). The CPS samples about 100,000 addresses annually and provides estimates for all states (101).

After several years of research to improve the measurement of uninsurance (83, 109) the Census Bureau revised the CPS' questions on health insurance coverage and income in 2014. The revised survey also added new questions on health insurance exchange participation, point-in-time coverage and employer-offers of coverage (110). Information on monthly insurance coverage data was also added and is currently being assessed for quality and disclosure concerns (108).

The CPS has not historically been used for longitudinal analysis due to the complexity of the files, but this analysis is supported by the survey design. Work by the Minnesota Population Center (MPC) has linked the CPS data over time to make longitudinal analysis easier (85). These linked data are provided through the Integrated Public Use Microdata Series (IPUMS) (51) and offer new opportunities to leverage the CPS content.

Survey of Income and Program Participation (SIPP-EHC)

The Survey of Income and Program Participation (SIPP-EHC) (96) collects longitudinal data on income and public program participation (e.g. Medicaid and Food Stamps) for the civilian non-institutionalized population using an event history calendar. The SIPP was re-

engineered for the 2014 panel to “reduce survey costs and respondent burden” (84, 97) eliminating the use of topical modules (104). Households are interviewed in person and by phone, once a year for four years. The 2014 SIPP-EHC panel began with a sample of about 53,000 households (104).

The SIPP-EHC has sample sizes designed to be representative for 20 states, but identifiers for all states are available on the public use files. The SIPP-EHC is designed specifically for longitudinal analysis, which is ideal for studying changes over time, but it has the longest lag for data release of any of the surveys discussed in this paper. The first wave from the 2014 panel (about coverage in 2013) was not released until early 2017.

Behavioral Risk Factor Surveillance System (BRFSS)

The Behavioral Risk Factor Surveillance System (BRFSS) (15) collects data regarding behavioral health and related risk factors for the adult civilian non-institutionalized population, is sponsored by the Centers for Disease Control and Prevention and is administered at the state level. In 2011, BRFSS updated its telephone-based sampling design to include cell phones (14). All states and the District of Columbia are required to use the survey’s core standardized questionnaire, which can be either fixed or rotated on a biannual basis. States have the option to include additional modules or state-specific questions.

A strength of the BRFSS is its flexibility for states to add their own content and be agile to changing policies. The flexibility of the BRFSS also creates challenges when working with the data. Comparisons can be difficult if content is asked of only a few states or removed from the core questionnaire.

National Health Interview Survey (NHIS)

The National Health Interview Survey (NHIS) (77) collects data on health status, access, utilization, and health behaviors of the civilian non-institutionalized population and is sponsored by the CDC’s National Center for Health Statistics (NCHS). The in-person survey is conducted continuously throughout the year with about 35,000 households sampled annually (69).

The 2014 and 2015 health insurance coverage estimates were published for all states and the District of Columbia. Due to funding constraints, the sample size was reduced in 2016 and insurance coverage estimates for all states are no longer available (17, 62). Researchers conducting analyses requiring state identifiers must go through the NCHS or Census Bureau’s Research Data Center (RDC) network.

NCHS has proposed redesigning the NHIS questionnaire in 2018 to reduce response burden and “to establish a long-term structure of ongoing and periodic topics” (76). The family level questions will be discontinued, with much of the content incorporated into the sample adult and sample child questions. With most questions asked of only the sample adult and sample child, the redesigned survey will reduce the sample size available for many questions.

Medical Expenditure Panel Survey – Household Component (MEPS-HC)

The Medical Expenditure Panel Survey – Household Component (MEPS-HC) (4) collects data on health, health care access, and expenditures and is sponsored by the Agency for Healthcare Research and Quality (AHRQ). The MEPS-HC is a longitudinal survey with participants interviewed in-person five times over a two-year period. Participants are selected from households included in the previous year of the NHIS. The sample includes about 13,500 families (5).

Information is also collected from each respondents' health care provider and supplements the household information. The MEPS-HC provides a breadth of health related information but is not state-representative. Only select estimates are released for the larger states with sufficient sample size. Researchers conducting analyses requiring state identifiers must go through the AHRQ Data Center or the Census/NCHS RDC network.

Harmonization efforts: NHIS and MEPS-HC

The Minnesota Population Center (MPC) has made significant efforts to harmonize federal health data that crosses years of data collection. The Integrated Health Interview Series (IHIS), funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, provides easily accessible integrated microdata for 5.5 million persons surveyed over 50 years, with more than 14,000 variables describing population health. The MPC is currently working on expanding the IHIS by including the MEPS-HC, the most complete source of information on population health care use and expenditures.

Linking NHIS and MEPS-HC

The MEPS household sample is drawn from the pool of respondents to the previous year's NHIS allowing each MEPS-HC annual release to be linked to the previous year's NHIS annual release beginning with the 1995 NHIS and the 1996 MEPS (59). These linked data files are currently only available through the AHRQ Data Center or the NCHS/CDC RDC network (3).

The NHIS-MEPS linkage allows researchers to access a variety of sociodemographic, health status, and household characteristics included in the NHIS but not the MEPS. For example, researchers have used these combined data to analyze the medical expenditures of immigrants in the U.S. (51, 89); the time and financial burden of caregiving to children with chronic conditions (112); and the effects of Medicaid eligibility on mental health services and out-of-pocket spending on mental health services (38).

3. Linking Administrative Records and Survey Data

Federal survey data may be linked with administrative data to better understand the health impact of policies but also to reduce respondent burden, improve accuracy, and reduce cost. We highlight data linkages that are currently maintained and updated for general use by extramural researchers. Some of these data are provided as public use files while others can be accessed only as part of the Census Bureau/NCHS RDC network.

NHIS and Death Records Longitudinal Mortality Files (LMF)

The NHIS for 1985–2009 is linked to longitudinal mortality data from the National Death Index (NDI) through Dec 31, 2011 (72) and was last updated in 2013 (105). A limited set of data are available through public use files with more data available as restricted files through the Census Bureau/NCHS RDC network (19). Restricted use files include detailed mortality information for all survey participants (including children) and information on age and cause of death (105). Published studies have used the linked data to evaluate the association between mortality and various health behaviors, conditions, and treatments, and compare mortality rates of different groups, adjusted for covariates (20).

NHIS and Centers for Medicare and Medicaid Services (CMS)

The 1994–2013 NHIS are linked to 1999–2013 Medicare enrollment and claims records collected by the Centers for Medicare and Medicaid Services (CMS) (70). In each of the last ten available NHIS samples, approximately 3,000–10,000 NHIS respondents are linked to Medicare administrative data (13–18% of linkage-eligible respondents) (33). Research files are available through the Census Bureau/NCHS RDC network (79).

Researchers have used these data to study hospitalization, readmission, and death among Medicare enrollees (20); to compare rates of self-reported diabetes in the NHIS with rates of diabetes identified in Medicare claims data (21); to study the relationship between moral hazard and the invasiveness of surgical procedures (57); and to assess NHIS measurement error of Medicare coverage (37).

NHIS and Social Security and Benefit History Data

The 1994–2005 NHIS are linked to the Old Age Survivors and Disability Insurance (OASDI) and Supplemental Security Income (SSI) records from the Social Security Administration (SSA) for respondents who agreed to provide their Social Security Number along with their name and date of birth (39, 73). Available data include benefits received from and payments to the SSA, eligibility for SSI disability benefits, and the SSA's determination of disability status for individuals receiving or applying for disability benefits. In each of the last ten available NHIS samples, 30,000–46,000 NHIS respondents are linked to SSA administrative data (44–62% of any given NHIS sample) (39). SSA-linked data are confidential and are available only through the Census Bureau/NCHS RDC network (75).

NHIS and Department of Housing and Urban Development (HUD)

The 1990–2012 NHIS are linked to HUD administrative data from 1999 through 2014 (56, 71). The administrative data include information on participation in public housing, type and timing of housing assistance received and housing structure (74). Approximately 1,300 to 2,600 respondents were linked in each of the last ten NHIS survey years (about 8–10% of linkage-eligible respondents) (66). HUD-linked data are confidential and are available only through the Census Bureau/NCHS RDC network (71).

4. New Aggregate Health Care Claims Data

The private sector has responded to the need for aggregate claims data to provide information on the trends in use and cost of health care services. Several firms serve as data aggregators that harmonize claims data across multiple payers creating large scale-Big Data for the purposes of data analytics and research. These data vendors vary by the number and type of participating health plans and by the accessibility of the data for health services research. We review four of the key data vendors and include information on sample coverage, sample size, geographic coverage and key variables.

Fair Health

Fair Health was established in 2009 as part of a legal settlement concerning New York State health insurance industry reimbursement practices (29). Fair Health receives data from 60 contributors including insurers and third party administrators and represents 150 million covered lives, approximately 75% of the privately insured population and 23.4% of national payments by privately insured patients. Costs for accessing the data are made on a case-by-case basis.

Researchers have used Fair Health data to study the effects of medical malpractice damage caps on provider reimbursement (34), market power and provider consolidation (54), the effects of occupational licensing on health care prices (53), the effects of health coverage mandates on provider reimbursement (35), and the opioid crisis among the privately insured (28).

MarketScan Research Databases

MarketScan Research Databases are a product of Truven Health Analytics, a subsidiary of IBM (13) that includes data from 150 employers, 21 commercial health plans and Medicare and Medicaid (22). The databases cover 230 million unique patients since 1995, and the most recent year includes 50 million covered lives (1). Researchers can analyze MarketScan data with Truven's proprietary analytic tools or on their own through licensing agreements.

Researchers have used MarketScan data to investigate the relationship between physician practice competition and physician service prices (7), the effects of managed care on angioplasty procedure prices (22), the impact of cost sharing on treatment adherence and outcomes among patients with diabetes (36), the relationship between incentive-based drug formularies and drug selection and spending on hypertension (52), the impact of hospital market consolidation on health care prices (63), and the causes of the 2009–2011 slowdown in health care spending (85).

Health Care Cost Institute Database

Health Care Cost Institute (HCCI) is a private non-profit established in 2011 to collect and disseminate health care cost and utilization data for Americans with private health insurance (45). The majority of its funding comes from its four data contributors: Aetna, Humana, Kaiser Permanente, and UnitedHealthcare (41), and its data covers 50 million individuals (43). The data cover 25% of the nonelderly population with employer-sponsored coverage

(43). Researchers can access HCCI data through its Data Enclave, a secure, virtual environment hosted by NORC at the University of Chicago that allows HCCI to maintain data security requirements (46).

Researchers have used HCCI data to examine the trends driving the growth in health spending among those with employer-sponsored coverage (50), the relationship between structural change in the health sector and health spending (23), the relationship between regional hospital prices and health spending (18), differences in reimbursement rates between Medicare Advantage and Medicare fee-for-service (6), and out-of-pocket spending on inpatient medical services among the non-elderly (2).

Optum Labs Data Warehouse

Optum Labs was founded as a partnership between Optum (a subsidiary of UnitedHealth Group, a public for-profit health insurer) and Mayo Clinic in 2013 and includes data from providers' electronic health record systems as well as from claims and enrollment systems of affiliated and non-affiliated health plans. The data cover 150 million lives, 19% of the population in commercial health plans, 19% of those in Medicare Advantage plans, 24% of those in Medicare Part D only plans, and 7% of the U.S. population with any health care utilization. So far, researchers have not published work using Optum Labs data to investigate questions related to health policy.

5. CONCLUSIONS

The U.S. federal survey resources are essential for monitoring trends in the nation's health, health insurance coverage and access to needed care, and health care spending. Yet, federal surveys face significant challenges. First from resource constraints and pressures to demonstrate their perceived utility in real time (16). For example, some federal surveys are more agile than others in responding to the need for timely data and adding survey content in response to shifts in the policy environment. For example, the NHIS significantly expanded its content in anticipation of ACA implementation and early release data from NCHS are the first indicators on health insurance coverage, released in the first quarter of the year reporting on prior year coverage (89).

There are also challenges due to the trends in survey research that include falling response rates, increase in non-response bias and changing modes of data collection. For example, in 1997 the NHIS achieved 92% household response (67) but fell to about 70% in 2015 (68). Surveys must be responsive. The ACS has gone a long way to meet respondent preferences by offering mail and on-line survey options as well as telephone and in-person follow-up for non-responders. While response rates are but one measure of quality; to date studies examining trends in bias between 1995 and 2015 are reassuringly stable (25).

These are also challenges to maintaining and assuring privacy as Americans signal a growing mistrust of government regarding privacy and confidentiality concerns (57). Leveraging administrative data is one option to reduce respondent burden, improve accuracy and reduce costs (98, 100). This is particularly appealing for sensitive topics, such as income that have high rates of missing data, which could be accessed through an administrative data

source. Again, this is not without challenges given issues of confidentiality and need for cooperation within agencies and across sectors (59).

Finally, some questions of interest to health services research require detailed data on use and costs of services—data that stretch the limits of self-report in surveys. This need has motivated the development of new frontiers in the data resource landscape such aggregated claims data from private sources. Health services researchers have much to discover and document about the strengths and limits of new data resources, while continuing to rely on available federal surveys and linked data resources. And federal agencies will need to continue to innovate to keep pace with the changes in survey methodology, trends in health and health care during a time of increased demands for data with limited resources.

LITERATURE CITED

1. Adamson, DM., Chang, S., Hanson, LG. Health Research Data for the Real World: the MarketScan Databases. Ann Arbor: Truven Health Anal; 2011 Jul. http://truvenhealth.com/portals/0/assets/PH_11238_0612_TEMP_MarketScan_WP_FINAL.pdf
2. Adrion ER, Ryan AM, Seltzer AC, Chen LM, Ayanian JZ, Nallamothu BK. Out-of-pocket spending for hospitalizations among nonelderly adults. JAMA Intern. Med. 2016; 9(176):1325–32.
3. AHRQ (Agency Healthc. Res. Qual.). Data File Types: Household Component-National Health Interview Survey Linked Files. AHRQ, Rockville, MD: https://meps.ahrq.gov/mepsweb/data_stats/more_info_download_data_files.jsp#hc-nhis
4. AHRQ (Agency Healthc. Res. Qual.). Medical Expenditure Panel Survey. AHRQ, Rockville, MD: <https://meps.ahrq.gov/mepsweb/>
5. AHRQ (Agency Healthc. Res. Qual.). Medical Expenditure Panel Survey: MEPS-HC Sample Sizes. AHRQ, Rockville, MD: 2015. https://meps.ahrq.gov/survey_comp/hc_sample_size.jsp
6. Baker LC, Bundorf MK, Devlin AM, Kessler DP. Medicare Advantage plans pay hospitals less than traditional Medicare pays. Health Aff. 2016; 8(35):1444–51.
7. Baker LC, Bundorf MK, Royalty AB, Levin Z. Physician practice competition and prices paid by private insurers for office visits. JAMA. 2014; 16(312):1653–62.
8. Barr, P., Lafave Grace, C. Thomson Reuters to sell Ann Arbor-based health care business to Veritas Capital. Crain News Serv. 2012 Apr 27. <http://www.craindetroit.com/article/20120423/FREE/120429974/thomson-reuters-to-sell-ann-arbor-based-health-care-business-to>
9. Bhaskar, R., Noon, J., O'Hara, B., Velkohff, V. CARRA Work. Pap. #2016-12. Washington, DC: Cent. Adm. Rec. Res. Appl., US Census Bur.; 2016 Dec 9. Medicare Coverage and Reporting: A Comparison of the Current Population Survey and Administrative Records. <https://www.census.gov/content/dam/Census/library/working-papers/2016/adrm/carra-wp-2016-12.pdf>
10. Bloomberg. New York: Bloomberg; <http://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=8936136>
11. Boudreaux MH, Thiede Call K, Turner J, Fried B, O'Hara B. Measurement error in public health insurance reporting in the American Community Survey: evidence from record linkage. Health Serv. Res. 2015; 50(6):1973–95. [PubMed: 25865628]
12. Borah, BJ. Optum Labs Overview. Rochester, MN: Mayo Clin.; 2016. <https://www.allianceforclinicaltrialsinoncology.org/main/cmsfile?cmsPath=/Public/Annual%20Meeting/files/Prevention-Optum%20Labs%20Overview.pdf>
13. Cao, J. IBM buying Truven for \$2.6 billion to amass more health data. Bloomberg. 2016 Feb 18. <https://www.bloomberg.com/news/articles/2016-02-18/ibm-to-acquire-truven-health-analytics-for-2-6-billion>
14. CDC (Cent. Dis. Control Prev.). Methodologic Changes in the Behavioral Risk Factor Surveillance System in 2011 and Potential Effects on Prevalence Estimates. Atlanta: CDC; 2016 Oct 31. <https://www.cdc.gov/surveillancepractice/reports/brfss/brfss.html>

15. CDC (Cent. Dis. Control Prev.). Behavioral Risk Factor Surveillance System. CDC, Atlanta: 2017 May 8. <https://www.cdc.gov/brfss/>
16. Census Proj. Home Page. 2017. <https://thecensusproject.org/>
17. Cohen, RA., Zammitti, EP., Martinez, ME. Health insurance coverage: Early release of estimates from the National Health Interview Survey, 2016. Hyattsville, MD: Natl. Cent. Health Stat.; 2017. <https://www.cdc.gov/nchs/data/nhis/earlyrelease/insur201705.pdf>
18. Cooper Z, Craig SV, Gaynor M, Van Reenen J. The price ain't right? Hospital prices and health spending on the privately insured. Work. Pap. No. w21815, Natl. Bur. Econ. Res. 2015
19. Data Link. Team. Comparative Analysis of the NHIS Public-use Linked Mortality Files: 2015 Public-use Data Release. Hyattsville, MD: Natl. Cent. Health Stat.; 2015 Feb. <https://www.cdc.gov/nchs/data/datalinkage/nhis-public-restr-2011lmf-2-3-15.pdf>
20. Data Link. Team. Linked Mortality Files Citation List. Hyattsville, MD: Natl. Cent. Health Stat.; 2016 Dec 16. https://www.cdc.gov/nchs/data/datalinkage/linked_mortality_files_citation_list_12_2016.pdf
21. Day, HR., Parker, JD. Natl. Health Stat. Rep. #84. Hyattsville, MD: Natl. Cent. Health Stat.; 2013 Nov 1. Self-report of Diabetes and Claims-based Identification of Diabetes among Medicare Beneficiaries. <https://www.cdc.gov/nchs/data/nhsr/nhsr069.pdf>
22. Dor A, Koroukian SM, Grossman M. Managed care discounting: evidence from the MarketScan database. *Inquiry*. 2004; 2(41):159–69.
23. Dranove D, Garthwaite C, Ody C. Health spending slowdown is mostly due to economic factors, not structural change in the health care sector. *Health Aff*. 2014; 8(33):1399–406.
24. Dunn A, Liebman E, Pack S, Shapiro AH. Medical care price indexes for patients with employer provided insurance: nationally representative estimates from MarketScan data. *Health Serv. Res*. 2013; 48(3):1173–90. [PubMed: 23088562]
25. Dutwin D, Buskirk T. Telephone sample surveys: Dearly beloved or nearly departed? Trends in survey errors in the age of declining response rates. Unpublished manuscript under peer review. 2017
26. Evans-Lacko SE, DosReis S, Kastelic E, Riley AW. Mental health service use before and after diagnosis of early-onset bipolar disorder. *J. Behav. Health Serv. Res*. 2011; 38(3):398–413. [PubMed: 20842462]
27. FAIRHealth. Overview of the FAIR Health Research Support Program Featuring the FAIR Health National Private Insurance Claims (FH NPIC®) Database. New York: FAIRHealth; 2014. <http://www.fairhealth.org/servlet/servlet.FileDownload?file=01560000000YXya>
28. FAIRHealth. The Opioid Crisis among the Privately Insured: The Opioid Abuse Epidemic as Documented in Private Claims Data. New York: FAIRHealth; 2016. <http://www.fairhealth.org/servlet/servlet.FileDownload?file=01532000001nwD2>
29. FAIRHealth. About FAIR Health. FAIRHealth; New York: 2017. <http://www.fairhealth.org/About-FH>
30. FAIRHealth. A Unique Program to Support Healthcare Research. FAIRHealth; New York: 2017. <http://www.fairhealth.org/Research>
31. FAIRHealth. FAQs. FAIRHealth; New York: 2017. <http://www.fairhealth.org/CorporateFAQ#>
32. FAIRHealth. Research FAQs. FAIRHealth; New York: 2017. <http://www.fairhealth.org/Research?sk=RESEARCH%20FAQS>
33. Fernandez, LE., Rastogi, S., Ennis, SR., Noon, JM. Evaluating Race and Hispanic Origin Responses of Medicaid Participants Using Census Data. April 20. CARRA Work. Pap. #2015-01. Washington, DC: Cent. Adm. Rec. Res. Appl., US Census Bur.; 2015. <https://www.census.gov/content/dam/Census/library/working-papers/2015/adrm/carra-wp-2015-01.pdf>
34. Friedson AI. Medical malpractice damage caps and provider reimbursement. *Health Econ*. 2017; 1(26):118–35.
35. Friedson A, Marier A. Mandated health insurance and provider reimbursement via private insurance: evidence from the Massachusetts health reform. *Health Serv. Res. Manag. Epidemiology*. 2017; (4):1–6.

36. Gibson TB, Song X, Alemayehu B, Wang SS, Waddell JL, Bouchard JR, Forma F. Cost sharing, adherence, and health outcomes in patients with diabetes. *Am. J. Manag. Care.* 2010; 8(16):589–600.
37. Gindi R, Cohen RA. Assessing measurement error in Medicare coverage from the National Health Interview Survey. *Medicare & Medicaid Research Review.* 2012; 2(2)
38. Golberstein E, Gonzales G. The effects of Medicaid eligibility on mental health services and out of pocket spending for mental health services. *Health Serv. Res.* 2015; 50(6):1734–50. [PubMed: 26445915]
39. Golden, C., Driscoll, AK., Simon, AE., Judson, DH., Miller, EA., Parker, JD. *Vital Health Stat. Vol. 1.* Hyattsville, MD: Natl. Cent. Health Stat.; 2015 Sep. Linkage of NCHS Population Health Surveys to Administrative Records from Social Security Administration and Centers for Medicare & Medicaid Services. https://www.cdc.gov/nchs/data/series/sr_01/sr01_058.pdf
40. Gorina, Y., Pratt, LA., Kramarow, EA., Elgaddal, N. *Natl. Health Stat. Rep. #84.* Hyattsville, MD: Natl. Cent. Health Stat.; 2015 Sep 28. Hospitalization, Readmission, and Death Experience of Noninstitutionalized Medicare Fee-for-service Beneficiaries Aged 65 and Over. <https://www.cdc.gov/nchs/data/nhsr/nhsr084.pdf>
41. HCCI (Health Care Cost Inst.). 2014 Health Care Cost and Utilization Report. Washington, DC: HCCI; 2015. <http://www.healthcostinstitute.org/files/2014%20HCCUR%2010.29.15.pdf>
42. HCCI (Health Care Cost Inst.). IRS Form 990: Return of Organization Exempt from Income Tax. Washington, DC: HCCI; 2015. <http://www.healthcostinstitute.org/wp-content/uploads/2017/01/Website-Version-HCCI-2015-Form-990.pdf>
43. HCCI (Health Care Cost Inst.). 2015 Health Care Cost and Utilization Report: Analytic Methodology v5.0. Washington, DC: HCCI; 2016. <http://www.healthcostinstitute.org/files/HCCI%202015%20Methodology%20v5.0%2011.22.16.pdf>
44. HCCI (Health Care Cost Inst.). What Makes HCCI Data Unique?. Washington, DC: HCCI; 2016 Jun. http://www.healthcostinstitute.org/wp-content/uploads/2016/11/How-Data-Unique-June-2016_0.pdf
45. HCCI (Health Care Cost Inst.). About HCCI. HCCI; Washington, DC: 2017. <http://www.healthcostinstitute.org/about-hcci/>
46. HCCI (Health Care Cost Inst.). About the HCCI Data Enclave. Washington, DC: HCCI; 2017 Mar. <http://www.healthcostinstitute.org/wp-content/uploads/2017/03/About-the-HCCI-Data-Enclave-fact-sheet-March-2017.pdf>
47. HCCI (Health Care Cost Inst.). Data Contributors. Washington, DC: HCCI; 2017. <http://www.healthcostinstitute.org/about-hcci/data-contributors/>
48. HCCI (Health Care Cost Inst.). Member Enrollment Fields. Washington, DC: HCCI; 2017. http://www.healthcostinstitute.org/wp-content/uploads/2017/01/HCCI_Data_Dictionary_2008-2014-1.10.17.pdf
49. Herrera, C., Hargraves, J., Stanton, G. Issue Brief #5. Washington, DC: Health Care Cost Inst; 2013. The Impact of the Mental Health Parity and Addiction Equity Act on Inpatient Admissions. <http://www.healthcostinstitute.org/files/HCCI-Mental-Health-Parity-Issue-Brief.pdf>
50. Herrera CN, Gaynor M, Newman D, Town RJ, Parente ST. Trends underlying employer-sponsored health insurance growth for Americans younger than age sixty-five. *Health Aff.* 2013; 10(32): 1715–22.
51. IPUMS CPS. Home Page. Minn. Popul. Cent.; Minneapolis: <https://cps.ipums.org/cps/>
52. Kamal-Bahl S, Briesacher B. How do incentive-based formularies influence drug selection and spending for hypertension? *Health Aff.* 2004; 1(23):227–36.
53. Kleiner MM, Marier A, Park KW, Wing C. Relaxing occupational licensing requirements: analyzing wages and prices for a medical service. *J. Law Econ.* 2016; 2(59):261–91.
54. Kleiner SA, White WD, Lyons S. Market power and provider consolidation in physician markets. *Int. J. Health Econ. Manag.* 2015; 1(15):99–126.
55. Klerman JA, Davern M, Thiede Call K, Lynch V, Ringel JD. Understanding the current Population Survey's insurance estimates and the Medicaid 'undercount'. *Health Aff.* 2009; 28(6):w991–1001.

56. Lloyd, PE., Helms, VE. NCHS-HUD Linked Data: Analytic Considerations and Guidelines. Hyattsville, MD: Natl. Cent. Health Stat; 2016 Jul. https://www.cdc.gov/nchs/data/datalinkage/nchs_hud_analytic_considerations.pdf
57. Loosveldt, G., Dominique, J. Defining and assessing survey climate. In: Wolf, C.Joye, D.Smith, TW., Fu, Y., editors. The SAGE Handbook of Survey Methodology. London: SAGE; 2016.
58. Margolis J, Hockenberry J, Grossman M, Chou SY. Moral hazard and less invasive medical treatment for coronary artery disease: the case of cigarette smoking. Work. Pap. No. w20373, Natl. Bur. Econ. Res. 2014
59. Miller PV. Is There a Future for Surveys? Public Opin. Q. 2017; S1(81):205–12.
60. Mohanty SA, Woolhandler S, Himmelstein DU, Pati S, Carrasquillo O, Bor DH. Health care expenditures of immigrants in the United States: a nationally representative analysis. Am. J. Public Health. 2005; 95(8):1431–8. [PubMed: 16043671]
61. Montalbano A, Rodean J, Kangas J, Lee B, Hall M. Urgent care and Emergency department visits in the pediatric Medicaid population. Pediatrics. 2016; 4(137)
62. Moriarity, C., Parsons, V. 2016 Sample Redesign of the National Health Interview Survey; 2015 National Conference on Health Statistics, Bethesda. 2015 Aug. p. 24-26.http://www.cdc.gov/nchs/ppt/nchs2015/Moriarity_Tuesday_Salon_E_BB2.pdf
63. Moriya AS, Vogt WB, Gaynor M. Hospital prices and market structure in the hospital and insurance industries. Health Econ. Policy. Law. 2010; 4(5):459–79.
64. Nasseh K, Vujicic M, Glick M. The relationship between periodontal interventions and healthcare costs and utilization: evidence from an integrated dental, medical, and pharmacy commercial claims database. Health Econ. 2017; 26(4):519–27. [PubMed: 26799518]
65. NCHS (Natl. Cent. Health Stat.). NCHS-CMS Medicare Linked Data: Match Rate Tables. Hyattsville, MD: NCHS; https://www.cdc.gov/nchs/data/datalinkage/cms_medicare_linked_data_match_rate_tables.pdf
66. NCHS (Natl. Cent. Health Stat.). NCHS-HUD Linked Data: Match Rate Table 1. Hyattsville, MD: Natl. Cent. Health Stat; https://www.cdc.gov/nchs/data/datalinkage/hud_table_1.pdf
67. NCHS (Natl. Cent. Health Stat.). 1997 National Health Interview Survey (NHIS) Public Use Data Release: NHIS Survey Description. Hyattsville, MD: NCHS; 2000.
68. NCHS (Natl. Cent. Health Stat.). 2015 National Health Interview Survey (NHIS) Public Use Data Release: Survey Description. Hyattsville, MD: NCHS; 2016. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2015/srvydesc.pdf
69. NCHS (Natl. Cent. Health Stat.). About the National Health Interview Survey. NCHS; Hyattsville, MD: 2016 Sep 12. https://www.cdc.gov/nchs/nhis/about_nhis.htm
70. NCHS (Natl. Cent. Health Stat.). NCHS Data Linked to CMS Medicare Enrollment and Claims Files. NCHS; Hyattsville, MD: 2016 Nov 29. <https://www.cdc.gov/nchs/data-linkage/medicare.htm>
71. NCHS (Natl. Cent. Health Stat.). NCHS Data Linked to HUD Housing Assistance Program Files. NCHS; Hyattsville, MD: 2016 Jul 7. <https://www.cdc.gov/nchs/data-linkage/hud.htm>
72. NCHS (Natl. Cent. Health Stat.). NCHS Data Linked to NDI Mortality Files. NCHS; Hyattsville, MD: 2016 Dec 21. <https://www.cdc.gov/nchs/data-linkage/mortality.htm>
73. NCHS (Natl. Cent. Health Stat.). NCHS Data Linked to SSA Social Security Benefit History Files. NCHS; Hyattsville, MD: 2016 Mar 29. <https://www.cdc.gov/nchs/data-linkage/ssa.htm>
74. NCHS (Natl. Cent. Health Stat.). Restricted-Use NCHS-HUD Data. NCHS; Hyattsville, MD: 2016 Jul 7. <https://www.cdc.gov/nchs/data-linkage/hud-restricted.htm>
75. NCHS (Natl. Cent. Health Stat.). Restricted-Use NCHS-SSA Data. NCHS; Hyattsville, MD: 2016 Mar 29. <https://www.cdc.gov/nchs/data-linkage/ssa-restricted.htm>
76. NCHS (Natl. Cent. Health Stat.). 2018 Questionnaire Redesign. NCHS; Hyattsville, MD: 2017 Feb 3. https://www.cdc.gov/nchs/nhis/2018_quest_redesign.htm
77. NCHS (Natl. Cent. Health Stat.). National Health Interview Survey. NCHS; Hyattsville, MD: 2017 May 22. <https://www.cdc.gov/nchs/nhis/>
78. NCHS (Natl. Cent. Health Stat.). Public-Use NCHS-CMS Medicare Feasibility Files. NCHS; Hyattsville, MD: 2017 Feb 2. <https://www.cdc.gov/nchs/data-linkage/medicare-feasibility.htm>

79. NCHS (Natl. Cent. Health Stat.). Restricted-Use NCHS-CMS Medicare Data. NCHS; Hyattsville, MD: 2017 Feb 8. <https://www.cdc.gov/nchs/data-linkage/medicare-restricted.htm>
80. Nicholson S. Research opportunities of a new private health insurance claims data set. *Health Manag. Policy Innov.* 2012; 1(1):37–41.
81. Noon, JM., Fernandez, LE., Porter, SR. CARRA Work. Pap. #2016-11. Washington, DC: Cent. Adm. Rec. Res. Appl., US Census Bur; 2016 Dec 9. Response Error and the Medicaid Undercount in the Current Population Survey. <https://www.census.gov/content/dam/Census/library/working-papers/2016/adrm/carra-wp-2016-11.pdf>
82. Optum. Real World Health Care Experiences. Eden Prairie, MN: Optum; 2015. https://cdn-aem.optum.com/content/dam/optum3/optum/en/resources/guides/5302_Data_Assets_Chart_Sheet_ISPOR.pdf
83. Pascale J. Modernizing a major federal government survey: a review of the redesign of the current population survey health insurance questions. *J. Off. Stat.* 2016; 32(2):461–86.
84. Planalp, C., Sonier, J., Turner, J. SHADAC Brief #41. Minneapolis: State Health Access Data Assist. Cent; 2014 May. Using Recent Revisions to Federal Surveys for Measuring the Effects of the Affordable Care Act. http://www.shadac.org/sites/default/files/publications/FINAL_SHADAC_Brief_41.pdf
85. Rivera Drew JA, Flood S, Warren JR. Making full use of the longitudinal design of the Current Population Survey: methods for linking records across 16 months. *J. Econ. Soc. Meas.* 2014; 39(3):121–44. [PubMed: 26113770]
86. Ryu AJ, Gibson TB, McKellar MR, Chernew ME. The slowdown in health care spending in 2009–11 reflected factors other than the weak economy and thus may persist. *Health Aff.* 2013; 32(5): 835–40.
87. Smith, WR. The 2011 National Household Survey—the Complete Statistical Story. Ottawa, ON: Stat. Can; 2015 Jun 4. <http://www.statcan.gc.ca/eng/blog-blogue/cs-sc/2011NHSstory>
88. Stapleton, M., Steiger, D. Cognitive Testing of the 2016 American Community Survey Content Test Items: Summary Report for Round 1 and Round 2 Interviews. Washington, DC: US Census Bur; 2015 Jan. https://www.census.gov/content/dam/Census/library/working-papers/2016/acs/2016_Westat_05.pdf
89. State Health Access Data Assist. Cent. SHADAC Brief #34. Minneapolis: State Health Access Data Assist. Cent; 2013 Jan. NHIS Questionnaire Changes Addressing the Patient Protection and Affordable Care Act. http://www.shadac.org/sites/default/files/publications/NHIS_ACA_Brief34.pdf
90. Stimpson JP, Wilson FA, Eschbach K. Trends in health care spending for immigrants in the United States. *Health Aff.* 2010; 29(3):544–50.
91. Turner, J. SHADAC Brief #47. Minneapolis: State Health Access Data Assist. Cent; 2016 Jun. New CPS Health Insurance Content: Data Release Timeline and Question Text. http://www.shadac.org/sites/default/files/publications/CPS_Redesign_final.pdf
92. US Census Bur. American Community Survey (ACS). US Census Bur.; Washington, DC: <https://www.census.gov/programs-surveys/acs/>
93. US Census Bur. American Community Survey (ACS): Content Test. US Census Bur.; Washington, DC: <https://www.census.gov/programs-surveys/acs/methodology/content-test.html>
94. US Census Bur. American Community Survey (ACS): Sample Size and Data Quality. US Census Bur.; Washington, DC: <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/>
95. US Census Bur. Current Population Survey (CPS). US Census Bur.; Washington, DC: <https://www.census.gov/programs-surveys/cps.html>
96. US Census Bur. Survey of Income and Program Participation. US Census Bur.; Washington, DC: <https://www.census.gov/sipp/>
97. US Census Bur. SIPP-EHC: 2011 and 2012 Field Test Evaluation. Washington, DC: US Census Bur; 2013.
98. US Census Bur A. Administrative Records and Third Party Data Use in the 2020 Census Working Group. Washington, DC: US Census Bur; 2014. https://www2.census.gov/cac/nac/reports/2014_2020census_wg.pdf

99. US Census Bur. American Community Survey Design and Methodology. Washington, DC: US Census Bur; 2014 Jan 30. https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf
100. US Census Bur. Evaluating the Availability and Suitability of External Data Sources for Use in the American Community Survey. Washington, DC: US Census Bur; 2015 Jul 24. https://www.census.gov/content/dam/Census/programs-surveys/acs/operations-and-administration/2015-16-survey-enhancements/External_Data_Sources.pdf
101. US Census Bur. Current Population Survey (CPS): Complete Technical Documentation. US Census Bur.; Washington, DC: 2016. <https://www.census.gov/programs-surveys/cps/technical-documentation/complete.html>
102. US Census Bur. Frequently Asked Questions about Accessing Data in the Data Linkage Infrastructure. Washington, DC: US Census Bur; 2016 Sep 8. <https://www2.census.gov/about/linkage/faqs.pdf>
103. US Census Bur. Source and Accuracy of Estimates for Income and Poverty in the United States: 2015 and Health Insurance Coverage in the United States: 2015. Washington, DC: US Census Bur; 2016. <https://www2.census.gov/library/publications/2016/demo/p60-257sa.pdf>
104. US Census Bur. Survey of Income and Program Participation: SIPP Introduction & History. US Census Bur.; Washington, DC: 2016 Feb 29. <https://www.census.gov/programs-surveys/sipp/about/sipp-introduction-history.html>
105. US Census Bur. Mandatory vs. Voluntary Methods. US Census Bur.; Washington, DC: 2017 Mar 16. American Community Survey (ACS). <https://www.census.gov/programs-surveys/acs/methodology/mandatory-voluntary-methods.html>
106. US Census Bur. Census Bureau Administrative Data Inventory. Washington, DC: US Census Bur; 2017 May 8. <https://www2.census.gov/about/linkage/data-file-inventory.pdf>
107. US Census Bur. Current Population Survey (CPS): Supplemental Surveys. US Census Bur.; Washington, DC: 2017 May 3. <https://www.census.gov/programs-surveys/cps/about/supplemental-surveys.html>
108. US Census Bur. Data Linkage Infrastructure. US Census Bur.; Washington, DC: 2017 Mar 9. <https://www.census.gov/about/adrm/linkage/guidance.html>
109. US Dep. Commer. & US Census Bur. Supporting Statement. Washington, DC: US Off. Manag. Budg; 2013 Feb 8. https://www.reginfo.gov/public/do/PRAViewDocument?ref_nbr=201211-0607-002
110. Vornovitsky, M., Day, JC. Health Insurance Measurement and Estimates. Expert Meeting on Income, Poverty, and Health Insurance. 2015. <https://www.census.gov/library/working-papers/2015/demo/2015-Income-Poverty-Health-Insurance-Experts.html>
111. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Crown WH. Optum Labs: building a novel node in the learning health care system. *Health Aff.* 2014; 33(7):1187–94.
112. Zan H, Scharff RL. The heterogeneity in financial and time burden of caregiving to children with chronic conditions. *Matern. Child Health J.* 2015; 19(3):615–25. [PubMed: 24951130]

Table 1

Health Related Content by Federal Survey^a

Agency	Census Bureau			CDC-States BRFSS ^b	CDC-NCHS NHIS	AHRQ MEPS-HC
	ACS 2015	CPS 2016	SIPP-EHC 2014 panel			
Health Insurance Coverage						
Uninsured rate	✓	✓	✓	✓	✓	✓
Coverage type	✓	✓	✓	✓	✓	✓
Coverage obtained through exchange		✓	✓	✓	✓	✓
Access to Care						
Usual source of care				✓	✓	✓
Told provider accepts insurance					✓	
No trouble finding a doctor					✓	
Affordability of Health Care						
Delayed medical care due to cost					✓	✓
Unmet need due to cost				✓	✓	✓
Changes made to medical drugs due to cost					✓	✓
Trouble paying medical bills					✓	✓
Out of pocket expenditures		✓	✓		✓	✓
Health Care Utilization						
Provider visit in past year			✓	✓	✓	✓
Mental health services					✓	✓
Emergency department visit in past year					✓	✓
Health Care Quality						
Adult cancer screenings (pap smears, colorectal cancer, mammograms)					✓	✓
Flu shot				✓	✓	✓
Smoking cessation counseling					✓	✓
Health Behaviors						

Agency	Census Bureau			CDC-States BRFSS ^b	CDC-NCHS NHIS	AHRQ MEPS-HC
	ACS 2015	CPS 2016	SIPP-EHC 2014 panel			
Body Mass Index (BMI)/obesity				✓	✓	✓
Binge drinking				✓	✓	
Smoking				✓	✓	✓
Health Outcomes						
Self-reported health status		✓	✓	✓	✓	✓
Activities limited due to mental or physical health difficulty	✓	✓	✓	✓	✓	✓
Chronic disease				✓	✓	✓
Physical Environment						
Neighborhood connectedness					✓	
Physical activity				✓	✓	
Geography						
State identifiers available	✓	✓	✓	✓	Only in RDC	Only in RDC
Sub-state identifiers available	✓	✓		✓		

Notes:

^aThe survey content is based on the questionnaire for the most current year of available data.

^bThe BRFSS includes content in the core questionnaire; states also have the option to add additional questions through optional modules.

All of the surveys include demographic and economic variables such as age, race/ethnicity, educational attainment, marital status, work status, and income.

Table 2

U.S. and State Sample Sizes by Federal Survey^a

	ACS		CPS		SIPP-EHC		BRFSS		MEPS-HC	
	2015	2016	2016	2016	2014 Panel, Wave 1	2015	2015	2015	2014	2014
U.S. Total	3,147,005	185,487	185,487	72,919	434,382	34,875				
Alabama	47,476	3,522	3,522	1,784	7,950	437				
Alaska	6,619	2,361	2,361	133	3,657	NA				
Arizona	67,014	3,358	3,358	2,299	7,946	540				
Arkansas	29,605	3,404	3,404	2,370	5,256	252 ^b				
California	374,943	18,254	18,254	6,399	12,601	6,089				
Colorado	53,570	2,226	2,226	960	13,537	410				
Connecticut	35,787	1,848	1,848	525	11,899	324				
Delaware	9,017	1,890	1,890	159	4,070	NA				
District of Columbia	6,610	2,974	2,974	122	3,994	NA				
Florida	194,548	8,230	8,230	2,987	9,739	2,017				
Georgia	97,854	4,175	4,175	2,402	4,678	1,155				
Hawaii	14,124	3,111	3,111	234	7,163	343 ^b				
Idaho	15,725	2,898	2,898	311	5,802	NA				
Illinois	126,642	5,441	5,441	3,728	5,289	1,115				
Indiana	66,045	3,070	3,070	2,572	6,067	578				
Iowa	31,900	1,945	1,945	556	6,227	272				
Kansas	28,774	2,201	2,201	495	23,236	350				
Kentucky	44,749	2,120	2,120	2,427	8,806	528				
Louisiana	43,892	4,341	4,341	2,472	4,716	402				
Maine	13,059	1,323	1,323	228	9,063	NA				
Maryland	59,332	2,349	2,349	974	12,598	685				
Massachusetts	68,785	3,517	3,517	1,000	9,294	651				
Michigan	98,008	3,807	3,807	2,586	8,935	931				
Minnesota	54,811	2,363	2,363	1,030	16,761	454				

	ACS 2015	CPS 2016	SIPP-EHC 2014 Panel, Wave 1	BRFS 2015	MEPS-HC 2014
Mississippi	29,600	3,586	1,878	6,035	366
Missouri	61,586	2,436	1,108	7,307	553
Montana	9,841	3,393	161	6,051	NA
Nebraska	19,089	2,346	254	17,561	NA
Nevada	26,988	2,398	342	2,926	NA
New Hampshire	13,378	2,197	152	7,022	NA
New Jersey	87,815	3,872	1,234	11,465	915
New Mexico	19,072	4,053	2,185	6,734	NA
New York	195,742	7,433	2,567	12,357	1,731
North Carolina	98,184	4,293	2,570	6,698	862
North Dakota	7,869	2,339	128	4,972	NA
Ohio	118,123	4,639	2,788	11,929	884
Oklahoma	37,251	2,913	736	6,943	288 ^b
Oregon	39,992	2,639	763	5,359	411
Pennsylvania	128,145	4,902	3,275	5,740	1,166
Rhode Island	10,563	1,445	174	6,206	NA
South Carolina	48,023	2,864	2,342	11,607	369
South Dakota	8,742	1,875	122	7,221	NA
Tennessee	65,549	3,259	2,411	5,979	600
Texas	259,224	11,735	4,676	14,697	3,332
Utah	29,290	3,080	533	11,401	425 ^b
Vermont	6,326	2,104	92	6,489	NA
Virginia	83,472	3,413	1,246	8,646	889
Washington	71,804	3,195	1,338	16,116	832
West Virginia	18,051	3,517	343	5,957	NA
Wisconsin	58,578	2,361	1,067	6,188	583
Wyoming	5,819	2,472	61	5,492	NA

Notes:

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

^aSample sizes are from the most current year of available public use data.

MEPS-HC state-level data are a special tabulation provided by the Agency for Healthcare Research and Quality.

^bStates do not have enough Primary Sampling Units to calculate direct variance estimates; a generalized variance function can be used.

NA states do not have sufficient sample size to calculate direct estimates.

The National Center for Health Statistics does not publish or provide state sample sizes for the NHIS because they are considered restricted data. The U.S. total NHIS sample size for 2015 is 103,789.

Table 3

Health Care Claims Databases

	Fair Health	MarketScan	Health Care Cost Institute Database	Optum Labs Data Warehouse
Organization that owns the data	Fair Health Inc. (private non-profit)	Truven Health Analytics (private for-profit)	Health Care Cost Institute (private non-profit)	Optum (subsidiary of UnitedHealth Group – public, for-profit)
Data contributors	60–70 private insurance carriers participating in the program (32, 80)	150 employers (1), 21 commercial health plans (1), Medicare and Medicaid (24).	Aetna, Humana, Kaiser Permanente and UnitedHealthcare (47)	Affiliated and non-affiliated commercial health plans, provider EMR/EHR systems (82)
Sample size	150 million covered lives, data gathered from about 15 billion claims. Represents estimated 23.4% of national payments by privately insured patients (32). Represents 75% of privately insured population (32).	230 million unique patients (since 1995), most recent year includes 50 million covered lives (1).	Claims related to 50 million unique people including individual, group, and Medicare Advantage members (44) (25.3% of nonelderly population with ESI)	150 million unique lives: 19% of US population in commercial health plans, 19% of those in Medicare Advantage plans, 24% of those in Medicare PDP plans, and 7% of U.S. population with any health care utilization (82)
Geographic coverage	Covers every locality in U.S. (32)	Wide geographic range, but disproportionately covers South (7). 10–12 unidentified states for Medicaid sample. (1)	Unknown	Relatively geographically representative, concentrated in the South and Midwest (82).
Variables of interest				
Geographic level	Geozip (first three numbers of ZIP code), can be aggregated to State/MSA level	Geozip (first three numbers of ZIP code), can be aggregated to State/MSA level (7)	ZIP Code of with pops greater than 1350 (48). Core Based Statistical Area (only Metro areas with 50,000+ populations included)	Not clear—at least available by Census Region
Race/ethnicity	Unknown	Only for Medicaid (61)	No	Yes
Age	Yes (DOB)	Yes (DOB) (1)	Yes (DOB) (48)	Yes
Other demographics	Some patient and provider information is optional.	Gender, aid category for Medicaid populations (blind/disabled, Medicare eligible), employment status, relationship of patient to beneficiary urban/rural status (64)	Gender, relationship to policyholder	Gender, sociodemographic characteristics (111) Race, income, education, assets, health risk assessment, mortality available via linked secondary data sources (48).
Inpatient	✓	✓	✓	✓
Outpatient	✓	✓	✓	✓
Pharmacy	✓	✓	✓	✓
Lab	✓	✓	✓ (41)	✓
Behavioral		✓ (26)	✓ (49)	✓
Dental	✓	✓		
Type of claim: fee charged vs. paid claim	All claims contain the fee billed by provider. About 50% of claims report “allowed charge” (80)	Claims represent the allowed amount paid by the plan (7).	Claims represent the allowed amount paid by the plan (48).	Claims represent actual paid amounts (12)

	Fair Health	MarketScan	Health Care Cost Institute Database	Optum Labs Data Warehouse
Run-out period	Database is updated twice yearly. Claims have a 3-month run-out (31).	Analysts can choose between “Early View data” with no minimum run-out, “Standard Updates” with 3 month minimum run-out, and “Annual File” with at least 6-month run-out (1).	Annual claims submitted at end of CY (44). Claims have a 5–6 month run-out period depending on payer (43).	Unknown

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript