



ASPE
ASSISTANT SECRETARY FOR
PLANNING AND EVALUATION



OFFICE OF THE SECRETARY
**PATIENT-CENTERED OUTCOMES
RESEARCH TRUST FUND**

REPORT

Medicare Data Linkages for Conducting Patient-Centered Outcomes Research on Economic Outcomes

Prepared for
The Office of the Assistant Secretary for Planning and Evaluation (ASPE)
at the U.S. Department of Health and Human Services

By
NORC at the University of Chicago
and
Washington University in St. Louis

September 2022



OFFICE OF THE ASSISTANT SECRETARY FOR PLANNING AND EVALUATION

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

THE OFFICE OF HEALTH POLICY

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community, in order to develop, analyze, coordinate and provide leadership on health policy issues for the Secretary. HP carries out this mission by conducting policy, economic and budget analyses, assisting in the development and review of regulations, assisting in the development and formulation of budgets and legislation, and assisting in survey design efforts, as well as conducting and coordinating research, evaluation, and information dissemination on issues relating to health policy.

OFFICE OF THE SECRETARY – PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) was established as part of the 2010 Patient Protection and Affordable Care Act and is charged to build data capacity for patient-centered outcomes research. Coordinated by ASPE on behalf of the Department, OS-PCORTF has funded a rich portfolio of projects to meet emerging U.S. Department of Health and Human Services policy priorities and fill gaps in data infrastructure to enhance capabilities to collect, link, and analyze data for patient-centered research. For more information visit <https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf>

This report was funded by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) under Contract Number HHSP233201500048II of the HHS Office of the Assistant Secretary for Planning and Evaluation (ASPE). The work was carried out by NORC at the University of Chicago, Washington University in St. Louis, and ASPE. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of HHS, ASPE, or NORC. Readers should not interpret any statement in this product as an official position of ASPE or of HHS.

Suggested Citation: Brown, D., Srinivasan, M., Zott, C., Wilson, K., Dullabh, P., and Smith, S. Medicare Data Linkages for Conducting Patient-Centered Outcomes Research on Economic Outcomes. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. September, 2022.

CONTRIBUTING AUTHORS

Derek Brown, PhD, Associate Professor, Washington University in St. Louis
Mithuna Srinivasan, PhD, Principal Research Scientist, NORC
Courtney Zott, MPH, Senior Research Associate II, NORC
Kala Wilson, MPA, Senior Research Associate II, NORC
Prashila Dullabh, MD, Vice President and Senior Fellow, NORC
Scott Smith, PhD, Director, ASPE

TECHNICAL EXPERT PANEL MEMBERS

Sandra L. Decker, Agency for Healthcare Research and Quality
Judith Dey, Office of the Assistant Secretary for Planning and Evaluation
Lindsey Enewold, National Cancer Institute
Meagan Khau, Centers for Medicare & Medicaid Services
Lauren Hersch Nicholas, University of Colorado
Kakoli Roy, Centers for Disease Control and Prevention
Elizabeth Shenkman, University of Florida
Todd Wagner, Stanford University
Hossein Zare, Johns Hopkins University

KEY INFORMANTS

Partha Bhattacharya, National Institute on Aging
Cathy Bradley, University of Colorado Denver
Michael Chernew, MedPAC and Harvard Medical School
Andrew Dick, RAND Corporation
Peter Groeneveld, University of Pennsylvania
Katherine Hempstead, Robert Wood Johnson Foundation
Laura Keohane, Vanderbilt University
Matthew Maciejewski, Duke University
Ciaran Phibbs, Health Economics Resource Center (HERC)

PROJECT OFFICERS AND PROJECT LEADERSHIP

Sara Wei (ASPE)
Susan Lumsden (ASPE)
Scott Smith (ASPE)
Prashila Dullabh (NORC)
Rina Dhopeswarkar (NORC)
Mithuna Srinivasan (NORC)



Table of Contents

Abstract	1
Introduction	2
Economic Outcomes Relevant for PCOR	3
Importance of Medicare Claims Linkages for PCOR Economic Outcomes.....	4
Purpose and Objectives of the White Paper	4
Methods	5
Identification of Federally Funded Data Sources Linked or Linkable to Medicare FFS Claims	5
Engagement with Subject Matter Experts	7
Findings	8
Identified Federally Funded Data Sources Linked or Linkable to Medicare FFS Claims.....	8
Degree of Capture of PCOR-Relevant Economic Outcomes in Identified Data Sources	10
Features of the Resulting Medicare FFS Claims-Based Linked Data	13
Strengths and Limitations of the Current Medicare-Based Linkages Infrastructure	15
Discussion.....	16
Considerations.....	17
Study Limitations.....	19
Conclusion	19
Appendix A: Economic Outcomes Relevant to PCOR Analysis of Health Care Interventions	20
Appendix B: Key Informant Discussion Questions	23
Appendix C: Federally funded Data Sources Currently LINKED OR Linkable to Medicare FFS Claims	26
References.....	35

Table of Exhibits

Exhibit 1. Search Terms for Identification of Data Sources	6
Exhibit 2. Fields in the Federal Data Inventory	7
Exhibit 3. Select Characteristics of Federally Funded Data Sources Linked or Linkable to Medicare Claims	8
Exhibit 4. SDOH Dimensions in Federally Funded Data Sources Linked or Linkable to Medicare Claims, by Source Type	10
Exhibit 5. Economic Outcome Domains in Federally Funded Data Sources Linked or Linkable to Medicare Claims	11
Exhibit 6. Stakeholder Perspectives Represented in Federally Funded Data Sources Linked or Linkable to Medicare Claims	12
Exhibit 7. Economic Outcome Measures in Federally Funded Data Sources Linked or Linkable to Medicare Claims	13
Exhibit 8. Rubric to Evaluate the Potential of Claims-Based Linkages for PCOR Economic Evaluations	14
Exhibit 9. Potential of Resulting Linkages for PCOR Economic Evaluations, by Source Type	15
Exhibit 10. Strengths and Weakness of Resulting Linkages	16

Abstract

Background. This review identifies federally funded data linkages that may facilitate patient-centered outcomes research (PCOR) on economic outcomes for Medicare fee-for-service (FFS) beneficiaries. The work builds on a previous review, which identified a variety of data sources for PCOR economic research and recommended that linked data sources be considered to expand the availability of economic outcomes. The review of linked data on the Medicare FFS population will be valuable to researchers seeking to produce new empirical evidence and will help identify gaps and barriers in patient-centered economic research for this population.

Objective. To facilitate PCOR on economic outcomes for Medicare FFS beneficiaries by identifying and assessing the characteristics of federally funded administrative and survey data sources linked or linkable to Medicare claims.

Methods. We conducted a stepwise scan during April-June 2022 of federally funded data sources linked to Medicare FFS claims and where the linked data are available to non-governmental researchers. First, we reviewed data sources identified in a prior ASPE white paper and shortlisted those that are linked to Medicare FFS claims. We then conducted a targeted scan to supplement this curated list. Information from identified data sources was abstracted into a spreadsheet inventory with fields related to characteristics of each data source, economic outcome domains, specific economic measures, stakeholder perspectives, and Medicare linkage details (available as an Appendix to this paper). The results were reviewed and refined by discussions with a technical expert panel, and during key informant interviews with researchers and data experts. After the inventory was finalized, we ranked each data source for its potential for patient-centered economic evaluations on a high/medium/low scale, based on the number of economic outcome domains, measures, and stakeholder perspectives captured in the data source.

Results. We identified 18 federally funded data sources for the inventory. Of these, 12 were survey data sources, and the remaining were administrative data sources. A majority were linked at the individual, beneficiary level and were collected or updated annually. All linked data sources provided some information on social determinants of health (SDOH). Specifically, all captured social context, 17 captured health care context, 14 economic context, 10 education, and seven physical infrastructure. All sources included some health equity-related factors, although there was wide variation. For PCOR economic outcomes, most (15) captured direct medical costs (beyond Medicare FFS payments), while five captured indirect costs (e.g., lost wages from absenteeism), and seven had data on direct non-medical costs (e.g., transportation). The patient and payer perspectives were represented in most linked data sources. Only three captured caregiver perspectives, and one captured the employer perspective. Eight data sources have “high” potential, seven have “medium” potential, and three have “low” potential for patient-centered economic outcomes.

Conclusion. The current landscape of federal and federally funded data sources provides a good range of PCOR economic outcomes when linked to Medicare FFS claims. Data linkages significantly enhance the availability of PCOR economic measures, perspectives, and domains. However, even with linked data, few capture direct non-medical or indirect costs. Additionally, smaller sample sizes and high access costs are among other limitations that inhibit researchers’ ability to perform comprehensive research on PCOR economic outcomes. Expanding linkages to include additional data sources, and reducing barriers to existing data sources, remain important objectives for increasing high-quality, patient-centered economic research for a wide range of stakeholders.

Introduction

Medicare is a uniquely large program covering 64 million beneficiaries across the nation.¹ Medicare beneficiaries include nearly all people of ages 65 and older, as well as many younger individuals with disabilities or end-stage renal disease. Medicare covers people of all races and ethnicities, who live in all 50 states and selected territories, in both urban and rural areas. The size and scope of Medicare make it a highly influential payer in the United States (U.S.) health care system, and its coverage and policies are often mirrored by private insurers. While the core Medicare (Parts A and B) is an entitlement program, Medicare offers optional benefits at additional cost, and many people purchase private insurance products to complement their Medicare coverage.²

Medicare beneficiaries and their families often face complex health care decisions. Aging often leads to more complicated health care conditions and treatments, and Medicare program options permit beneficiaries to make annual decisions about their health insurance coverage options. Centers for Medicare & Medicaid Services (CMS) quality ratings, the Medicare plan finder, and other tools can help beneficiaries to navigate certain choices. Choices for medication, treatment modalities, chronic disease care, and disease management are challenging and may have substantial health and economic consequences for patients, families, caregivers, and providers. Improved data to support patient-centered decision-making could help to improve the health and financial well-being of Medicare beneficiaries and their families. To date, such information remains limited, although data generated by the Medicare system present a rich potential option for improving research on patient-centered outcomes. This paper inventories and evaluates linkages of Medicare claims data for conducting patient-centered outcomes research (PCOR) on economic outcomes.

Medicare fee-for-service (FFS) claims data hold several advantages for researchers seeking to study the health and health-related economic outcomes of Medicare beneficiaries.^{3,4,5} Generated from Medicare billing records, Medicare FFS claims data (hereafter referred to as “claims”) capture key pieces of information about the types of health care received by a patient; the resources and payments associated with that care; associated health care conditions; and basic patient demographics.⁶ FFS claims are reviewed and validated by CMS before payment occurs, improving data quality. Medicare claims are updated continuously and include multiple file types with an array of providers and benefit information. More than 40 percent of Medicare beneficiaries are now enrolled in Medicare Advantage plans; such plans submit encounter data to CMS that can also be used for research. Medicare data have been used extensively to study certain types of cost outcomes.^{7,8} At the patient level, several types of Medicare claims capture information about Medicare’s costs of treating a specific condition or reimbursement for selected prescription drugs. Some information about patient costs is also available, including coinsurance and copayments within the Medicare program.^a

There are also some disadvantages to using Medicare claims for health services research. They contain limited clinical information and omit details such as laboratory results or electronic health records. Medicare claims data also do not capture non-covered services, informal caregiving, supplemental insurance payments, or over-the-counter medications. Payment and premium information is also largely unavailable for persons enrolled in Medicare managed care plans (Part C or Medicare Advantage). Both FFS and Medicare Advantage data contain

^a For many beneficiaries, however, FFS claims provide only a partial picture of their coinsurance and copayments, owing to widespread purchase of supplemental insurance plans (“Medigap”), dual eligibility for Medicaid, or enrollment in Medicare managed care (Part C or Medicare Advantage). The Medicare Current Beneficiary Survey (MCBS) includes some additional data on patient costs and other information for a sample of beneficiaries.

utilization details on specific health services: diagnoses, procedures, dates, quantities, type of service, provider or facility identifiers, and basic patient demographics.⁹ FFS data also include charged, allowed, and paid amounts, and data on coinsurance and deductibles. A complete patient-centered perspective on economic costs includes a wide range of measures (described in the following subsection), many of which are not available in Medicare claims. To help address this, data linkages may enable researchers to add economic outcomes from multiple data sources, improving the responsiveness, relevancy, and breadth of patient-centered economic research that may be facilitated with Medicare claims data. This improved research may help patients, families, and clinicians to make more informed, patient-centered decisions about health care treatments, and may help payers and policymakers to identify opportunities for improving population health and patient outcomes.

Economic Outcomes Relevant for PCOR

Economic outcomes may be directly or indirectly associated with health and health care and include a wide range of costs and financial impacts.¹⁰ Living with a health condition, seeking medical care, providing medical treatments, and managing care all involve many different types of economic costs.⁵ The Patient Protection and Affordable Care Act (ACA) of 2010 established the Patient-Centered Outcomes Research (PCOR) Trust Fund to support activities to promote research on the outcomes and effectiveness of health care relevant to generate evidence for decision-making by patients, providers, and policymakers.¹¹ When the PCOR Trust Fund was reauthorized by Congress in 2019, the scope of PCOR studies was expanded to include economic and cost burdens of health care: “...clinical and patient-centered outcomes shall include the potential burdens and economic impacts of the utilization of medical treatments, items, and services on different stakeholders and decision-makers respectively. These potential burdens and economic impacts include medical out-of-pocket costs, including health plan benefit and formulary design, nonmedical costs to the patient and family, including caregiving, effects on future costs of care, workplace productivity and absenteeism, and healthcare utilization.”

In a [previous report](#), three domains of PCOR economic outcomes were defined, and the relevance of each of these domains to different stakeholder groups was discussed.¹² The three domains were: (1) direct medical costs (those tied to the receipt of specific health care services for a patient); (2) direct non-medical costs (those associated with health care or illness, but not part of the purchase of medical services); and (3) indirect costs (the value of resources that are lost as a result of incurring illness or health care treatments).¹³ For the older adult Medicare population, certain types of costs (e.g., informal caregiving, absenteeism) may appear to affect patients differently than working age adult populations, although indirect and other costs may be borne by family members, and distributional differences will vary by health care condition and other factors.

Patient-centered care decisions may be informed by economic outcomes, but non-economic factors may also be considered and may play a large role in certain situations and types of care. Economic costs and benefits are not the only factors that enter patient or family member decisions. Additionally, researchers may also include intangible costs, such as quality of life or expected length of life under economic impacts relevant to PCOR. However, valuation of such intangible costs is controversial.¹⁴ Furthermore, the expanded definition of PCOR in the 2019 reauthorization does not explicitly list intangible costs, so they were excluded from consideration in the previous ASPE report and from this report.

Importance of Medicare Claims Linkages for PCOR Economic Outcomes

The same ASPE report noted above also reviewed 29 federally funded data sources for their content on PCOR economic outcomes. Many of the sources were found to contain data on direct medical costs, but relatively fewer contained measures in the other two outcome domains. Linking data sources to connect the strengths and features of different datasets was identified as a key strategy for improving the PCOR economic outcomes landscape in the future.

In this regard, Medicare claims-based linkages for PCOR economic research are of high interest for several unique reasons. First, CMS has an established research data dissemination program that offers technical support services to assist researcher access and analysis of Medicare data, via the Research Data Assistance Center (ResDAC)^b, the Chronic Conditions Data Warehouse (CCW)^c, and the CCW Virtual Research Data Center (VRDC).^d It may also be easier for researchers to link other data sources to Medicare (including other CMS data such as Part D and Medicaid or other sources of data like surveys or registries) to support CMS-approved research goals, than some other types of administrative data (e.g., selected state data, proprietary provider or payer data). Second, not only is Medicare a major payer as described above, but relative to all other major payers, it also is projected that Medicare will experience the fastest growth rate of spending (7.6 percent per year over 2019-28).¹⁵ And finally, the Medicare population covers, on average, an older population with more substantial health care needs and higher prevalence of chronic conditions, so average per-patient medical costs in Medicare are greater than those of working age adults or children.¹⁵

Purpose and Objectives of the White Paper

This paper reviews Medicare claims-based data linkages for conducting PCOR on economic outcomes. The paper focuses on traditional, fee-for-service (FFS) Medicare “claims,” defined as Parts A and B coverage, not Medicare Advantage (Part C). We also assume that linked Medicare FFS claims data are based on the CMS research identifiable files (RIF) format. Unlike Parts A and B, Medicare Advantage data do not capture direct medical costs or most of the PCOR economic outcomes of interest that were described above. Medicare Advantage claims have only recently been released to the research community,¹⁶ and the research with these data is still nascent, compared to more than two decades for FFS claims.

We acknowledge that Part C data will be increasingly useful for future health services research, particularly as enrollment in Part C continues to grow. Enrollment in the Medicare Advantage program has been increasing since 2009, particularly among Black and Hispanic beneficiaries.¹⁷ Federally housed Medicare Advantage encounter data are useful for measuring health care utilization but do not contain direct estimates of payer costs or other economic outcomes. Researchers must therefore use other sources on plan coverage and cost-sharing requirements to derive cost estimates—for example, Plan Finder can be used to obtain Medicare Advantage beneficiaries’ monthly premiums, deductibles (for both medical services and prescription drugs), in-network and

^b See <https://resdac.org/>.

^c See <https://www2.ccwdata.org/>.

^d See <https://resdac.org/cms-virtual-research-data-center-vrdc>.

out-of-network copays (e.g., diabetes supplies and services, emergency care, specialist office visit) for a variety of services, and maximum out-of-pocket amounts.¹⁸

Two other exceptions to the definition of Medicare used in this paper are worth noting. Part D Prescription Drug Event (PDE) data provide summary information on prescription drug costs and payments for those Medicare beneficiaries who are voluntarily enrolled in a Part D plan (but not individual drug claims), and these data can be linked to Medicare FFS claims. Similarly, comprehensive, national-level data on health care utilization and expenditures are available for those individuals enrolled in Medicaid (which can support research on “dual eligibles”). While Part D data and Medicaid claims are included in this paper’s data inventory as auxiliary data sources that are linkable to Medicare FFS claims, they were excluded from the definition of Medicare FFS claims.

This paper has these key objectives:

1. To guide researchers on data linkages between Medicare FFS claims data and other relevant federally funded administrative and survey data sources that capture information on one or more of the PCOR economic outcome domains;
2. To identify the specific economic impacts that are measurable with each data linkage; and
3. To illustrate gaps that remain for PCOR economic outcomes for the Medicare population to inform priorities into additional Medicare data linkage efforts.

Methods

Our approach to compiling an inventory of federal and federally funded survey and administrative data sources (that have one or more economic outcomes and can be linked to Medicare claims) comprised these steps: (1) first, we identified a list of such data sources relevant to PCOR economic evaluations by building on the inventory from the previous published report; and (2) second, we sought feedback on these data sources from subject matter experts via a technical expert panel (TEP) and key informant interviews. Below we describe each of these steps in more detail.

It is important to note that economic outcomes considered for this study comprise those that were identified in the previous report. As explained earlier, these outcomes fall into three domains of direct medical costs, direct non-medical costs, and indirect costs. [Appendix A](#) shows specific measures that fall under each of these three domains, mapped to four stakeholder perspectives: the patient, caregiver, employer, and payer.

Identification of Federally Funded Data Sources Linked or Linkable to Medicare FFS Claims

Data sources eligible for inclusion in this study had to meet these inclusion criteria:

- Be federally funded;
- Capture at least one of the three economic outcome domains relevant for PCOR;
- Be linked or linkable to Medicare claims at the individual level; and
- Be a source where the resulting linkage with Medicare claims is available to external researchers.

We followed a stepwise approach during April-June 2022 to efficiently identify federal and federally funded data sources relevant for this study, which built on our existing knowledge of data sources. First, we reviewed the federal and federally funded data sources that were included in the previous report. While all those previously identified data sources contain one or more economic outcomes, not all are linked or linkable to Medicare claims. Therefore, we selected a subset of those previously identified data sources that are linked or linkable to Medicare claims. As a check, for those data sources that did not get included via this step, we did a web search to confirm that they indeed are not linked or linkable to Medicare claims.

In the second step, we reviewed the website of the Medicare and Medicaid Resource Information Center (MedRIC)¹⁹ to capture data sources not already identified via the prior step, and that meet the four inclusion criteria listed above. Next, we reviewed the websites of the federal agencies whose data sources were included via the above steps, to ensure that all their relevant data sources were captured. Finally, we performed a targeted internet search for additional data sources, as well as for relevant peer-reviewed literature that represents reviews or compilations of data sources. We relied on PubMed and Google Scholar to identify peer-reviewed sources published in English. Search terms for the scan of data sources and peer-reviewed literature are shown in [Exhibit 1](#), and we used Boolean operators to combine these search terms.

Exhibit 1. Search Terms for Identification of Data Sources

Stakeholder-Related	Economic Outcome-Related	Other
<ul style="list-style-type: none"> ■ Patient ■ Caregiver ■ Employer ■ Payer 	<ul style="list-style-type: none"> ■ Cost of illness ■ Burden of illness ■ Indirect cost ■ Productivity ■ Absenteeism ■ Presenteeism ■ Economic impact ■ Economic assessment ■ Economic burden ■ Direct medical cost ■ Direct non-medical cost ■ Indirect cost 	<ul style="list-style-type: none"> ■ Federal ■ Federally funded ■ Data ■ Medicare ■ Claims

Once the list of data sources was identified, relevant information was abstracted from each of them into a spreadsheet inventory. [Exhibit 2](#) lists the fields included in the inventory, organized by category. Fields have been grouped into five categories related to characteristics (e.g., the periodicity of data collection, and source of data), data access (whether the data are publicly available, and the purchase cost), size and scope (e.g., years of data available), economic outcomes (both the three domains of outcomes and available measures in each domain), and linkage with Medicare claims (e.g., available time frame for linked data).

Exhibit 2. Fields in the Federal Data Inventory

Fields Related to Characteristics	Fields Related to Data Access	Fields Related to Size and Scope	Fields Related to Economic Outcomes	Fields Related to Linkage with Medicare Claims
<ul style="list-style-type: none"> ■ Data source name ■ Data source acronym ■ Data source steward ■ Data source URL ■ Periodicity of data collection (e.g., annual) ■ Source of data (administrative or survey) ■ Administrative data type (e.g., hospital encounter records; NA for survey data) ■ Lowest level of aggregation (e.g., individual) ■ Length of observation (e.g., longitudinal) 	<ul style="list-style-type: none"> ■ Whether the data are publicly available or not ■ If publicly available, whether additional restricted-use data are available (NA for data sources not publicly available) ■ Whether there are costs to acquire data 	<ul style="list-style-type: none"> ■ Years of data available ■ Latest available sample size ■ If survey data, latest available response rate (NA for administrative data) ■ Population scope of data (e.g., U.S. civilian non-institutionalized population) ■ Whether the data are nationally representative/have national coverage or not ■ Lowest level of geography available 	<ul style="list-style-type: none"> ■ Available economic outcome categories (e.g., direct medical costs) ■ Available economic outcome measures, by category (e.g., out-of-pocket health care costs for direct medical costs) ■ Economic perspectives represented (e.g., patients) ■ Identifiable SDOH categories* ■ Whether demographics data are available for disparities/equity analysis** ■ Facility identifiers and information 	<ul style="list-style-type: none"> ■ Available time frame for linked data ■ How outside researchers can access the linked data ■ Rating of linked data for PCOR economic evaluations***

*This field was populated on the basis of five social determinants of health (SDOH) categories that were defined based on AHRQ’s framework (<https://www.ahrq.gov/sdoh/about.html>).

**In addition to age and sex, other demographic characteristics that were used to populate this field were derived from the *Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*.

***Linked data were rated as high, medium, or low. A “high” rating was assigned when multiple economic outcome domains and multiple stakeholder perspectives are available in a linkage; a “medium” rating was assigned when there is at least one element of multiplicity (e.g., multiple economic outcome domains or multiple perspectives or multiple outcomes within a given domain); all other data sources were assigned a “low” rating.

Engagement with Subject Matter Experts

We solicited subject matter expert feedback on the list of federal and federally funded data sources using two channels. First, we convened a nine-member TEP comprised of federal and non-federal experts. We included a diverse array of experts in PCOR, health economics and evaluation, health economic measurement, and health equity; federal members represented various agencies. We met with the TEP three times between May and September 2022 to gather feedback on data sources, as well as on a draft of this white paper.

We also conducted nine individual semi-structured interviews with key informants representing researchers, data stewards, and Medicare policymakers. Questions in the interviews centered around identifying data sources that are currently linked or linkable to Medicare claims, discussing potential linkages that will become available to researchers in the future, identifying gaps in the current federal data infrastructure that may pose barriers to robust and rigorous PCOR economic evaluations, and prioritizing other white paper topics related to

data relevant for PCOR economic evaluations. [Appendix B](#) contains the discussion guide used for these interviews.

Findings

Below we describe the key characteristics of identified data sources that are linked or linkable to Medicare claims, the economic outcomes available in each, features of the resulting Medicare claims-based linked data, and relative strengths and limitations. Summarized frequencies are presented for select characteristics in exhibits, and an abbreviated version of the data inventory underlying these summaries is available in [Appendix C](#), which can be used to identify the specific data sources being referenced. The full version of the data inventory is available on <https://aspe.hhs.gov/>.

Identified Federally Funded Data Sources Linked or Linkable to Medicare FFS Claims

We identified 18 federally funded data sources for inclusion in the inventory (not counting Medicare claims). [Exhibit 3](#) shows select characteristics of these data sources. Notably, survey data sources (12) are much more prevalent than administrative data sources (6). Survey data sources capture a wider range of social determinants of health (SDOH) dimensions and characteristics relevant for health equity than do administrative data sources (see [Exhibit 4](#)), indicative of their ability to enhance claims data.

Exhibit 3. Select Characteristics of Federally Funded Data Sources Linked or Linkable to Medicare Claims

Characteristic	Number of Data Sources	% of Data Sources ¹
Data Source Steward		
Centers for Disease Control and Prevention/ National Center for Health Statistics (CDC/NCHS)*	5	28%
CMS*	4	22%
National Institutes of Health (NIH)*	6	33%
University of Michigan - Ann Arbor	1	6%
U.S. Department of Veterans Affairs	1	6%
Agency for Healthcare Research and Quality (AHRQ)	1	6%
Source of Data		
Survey	12	67%
Administrative	6	33%
Lowest Level of Aggregation²		
Individual	17	94%
Encounter/claim	3	17%
Length of Observation³		
Panel/longitudinal	12	67%
Cross-sectional	7	39%
Periodicity of Data Collection⁴		
Annual	14	78%
Biennial	3	17%
Other ⁵	2	11%

Characteristic	Number of Data Sources	% of Data Sources ¹
Observable Social Determinants of Health (SDOH) Domains⁶		
Social context	18	100%
Economic context	14	78%
Education	10	56%
Physical infrastructure	7	39%
Health care context	17	94%
Observable Health Equity-Related Factors		
Age	18	100%
Sex	18	100%
Race/ethnicity**	14	78%
Income or income status**	12	67%
Urban-rural status**	16	89%
Disability status**	13	72%
Religious affiliation**	2	11%
LGBTQ+ status**	3	17%

Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

¹ Totals may not add up to 100 percent due to rounding or non-mutual exclusivity of categories.

² These categories are not mutually exclusive. Medicaid and Part D claims files include both individual-level enrollment data and encounter/claim-level data.

³ These categories are not mutually exclusive. The National Study of Caregiving has historically been a source of cross-sectional data. However, longitudinal data collection began in 2017, and will continue to be implemented going forward. Therefore, this data source was counted as both cross-sectional and longitudinal data.

⁴ These categories are not mutually exclusive. The Panel Study of Income Dynamics provided annual data from 1968-1997, and then changed to biennial after 1997. Therefore, this data source was counted as both annual and biennial data.

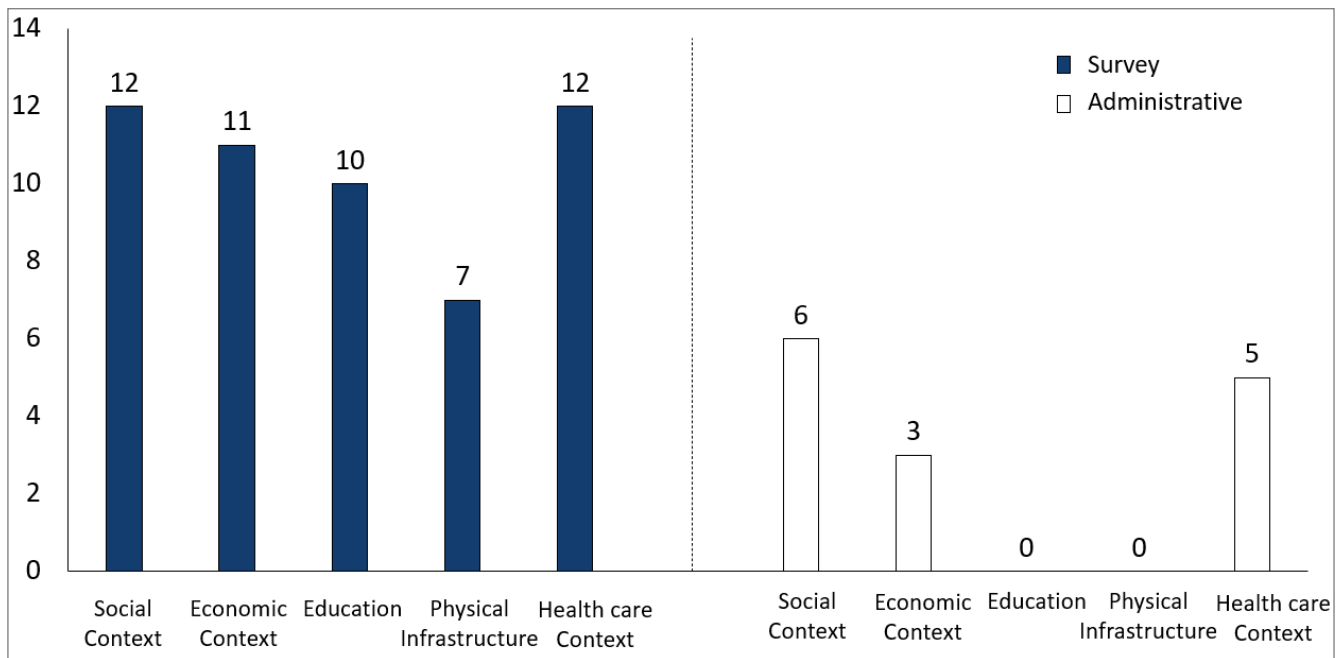
⁵ Values in "Other" include: (1) every five years (National Long Term Care Survey) and (2) 1973-74, 1977, 1985, 1995, 1997, 1999, and 2004 (National Nursing Home Survey)

⁶ These are based on AHRQ's framework (<https://www.ahrq.gov/sdoh/about.html>).

* Represents agencies at HHS.

**These characteristics have been identified as priority populations in the *Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*.

Exhibit 4. SDOH Dimensions in Federally Funded Data Sources Linked or Linkable to Medicare Claims, by Source Type



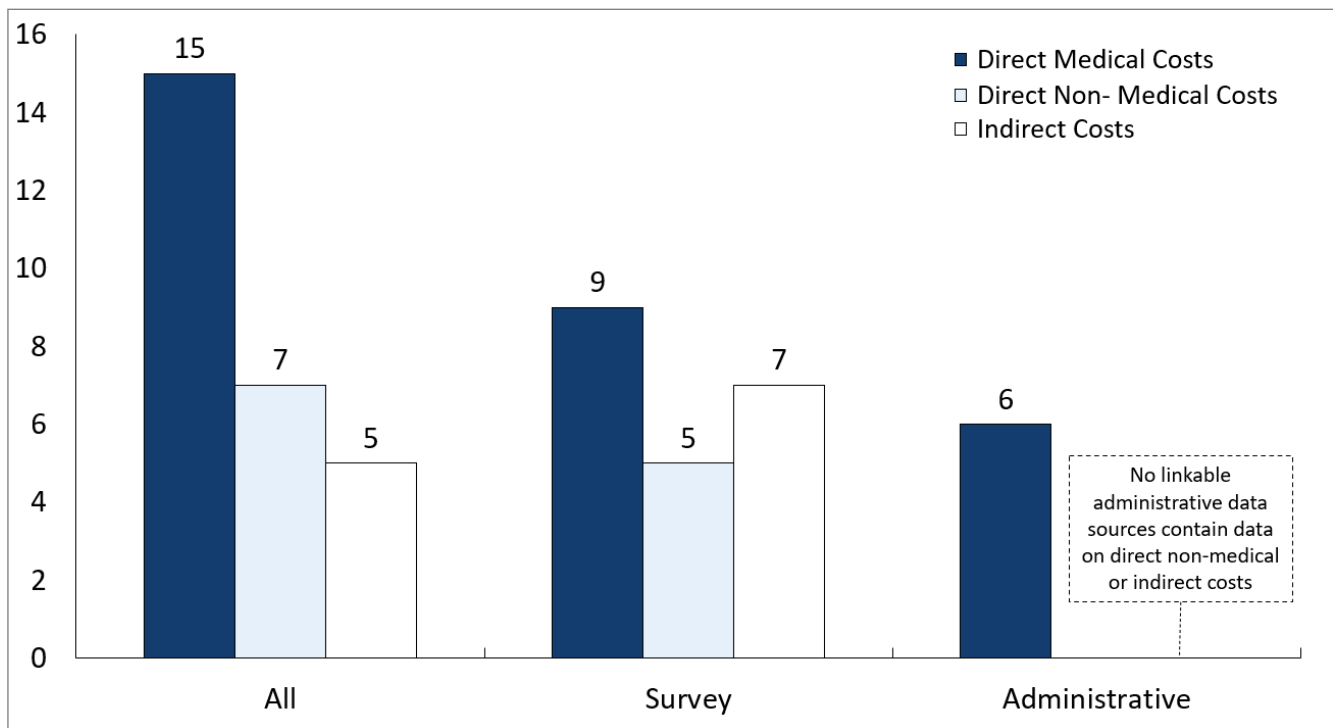
Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

Degree of Capture of PCOR-Relevant Economic Outcomes in Identified Data Sources

Exhibits 5 to 7 provide information on the degree of capture of PCOR-relevant economic outcomes of the data sources included in the inventory. The economic outcome categories and measures are not mutually exclusive across data sources; several capture more than one outcome category or measure.

Within Medicare claims, only direct medical costs are available, specifically, paid/reimbursed amounts. Among the auxiliary data sources linked or linkable to Medicare claims as well, most (15) contain information on direct medical costs (see [Exhibit 5](#)). Two data sources contain information on all three economic outcome domains—the Panel Study of Income Dynamics (PSID)²⁰ and the Medical Expenditure Panel Survey (MEPS)²¹—but there are relatively few measures within each domain. Similar to information on SDOH dimensions, survey data sources capture a wider range of economic outcomes than administrative data sources (see [Exhibit 5](#)). None of the identified administrative data sources contain information on direct non-medical or indirect costs. Additionally, despite being better sources for direct medical costs, key informants noted that administrative data sources lack information on out-of-pocket costs for beneficiaries with supplemental coverage (e.g., Medigap), preventing understanding of the true amount patients pay for individual encounters.

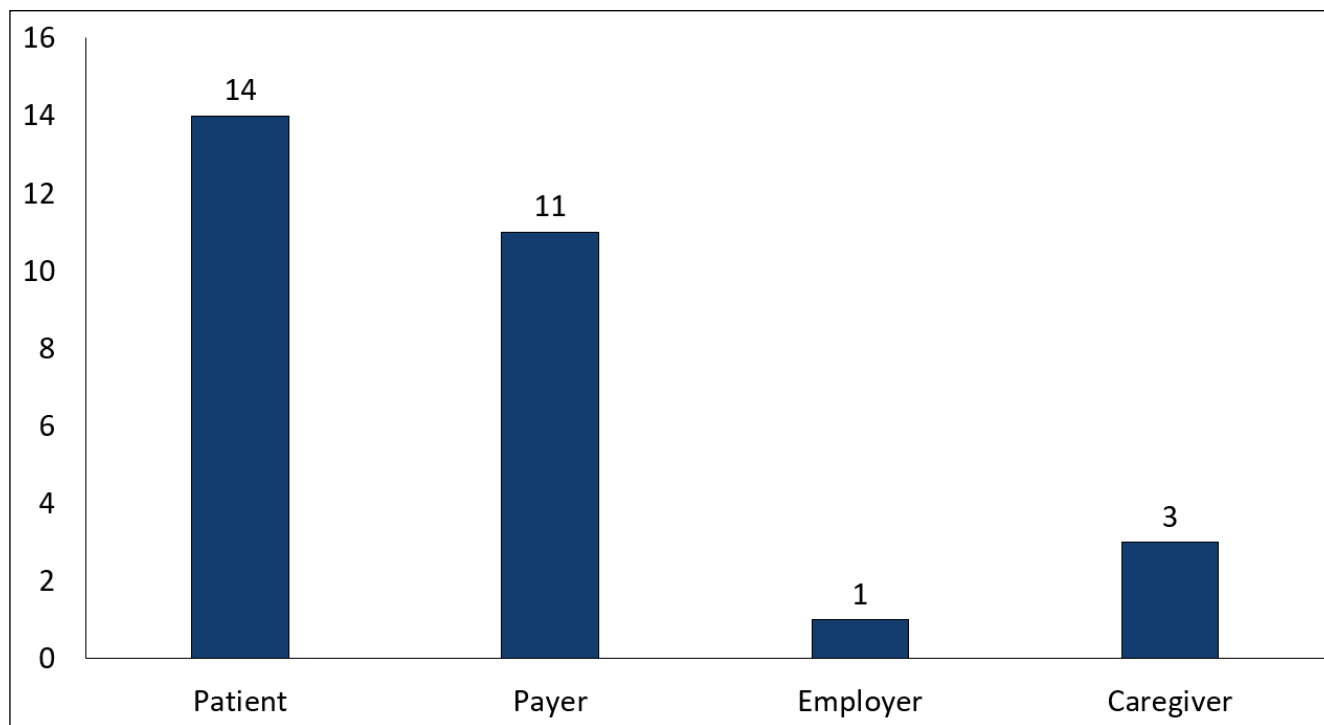
Exhibit 5. Economic Outcome Domains in Federally Funded Data Sources Linked or Linkable to Medicare Claims



Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

We also assessed the stakeholder perspectives—patients, payers, employers, and caregivers—who represent the entities most likely to use economic information—in the data sources for decision-making. As shown in [Exhibit 6](#), patient perspectives are most commonly available (in 14 data sources).

Exhibit 6. Stakeholder Perspectives Represented in Federally Funded Data Sources Linked or Linkable to Medicare Claims



Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

We also examined the specific outcome measures in each of the three domains that are observable in the 18 identified data sources. As seen in [Exhibit 7](#), out-of-pocket costs and paid/reimbursed amounts are the most commonly available measures both within the direct medical cost category and across all measures, available in 11 and 10 data sources respectively. Though direct non-medical and indirect cost measures are not widely available, the value of informal caregiving is most common in the former (in three data sources), and the value of absenteeism is most common in the latter (in five data sources).

Comparing the 18 data sources, we see that the National Study of Caregiving (NSOC),²² PSID, and MEPS are the most promising for PCOR economic evaluations when linked to Medicare data. The resulting claims-based linkages with these three data sources provide the most information on economic outcomes and stakeholder perspectives, enabling researchers to conduct more holistic assessments of the costs of care. Linked NSOC-Medicare data provide the richest source of information on both caregiver and patient perspectives on late-life care, covering all economic outcome domains and 10 economic measures: the cost of informal caregiving, transportation, special food, time spent seeking health care, home modifications, housekeeping, absenteeism, inability to work, out-of-pocket health care costs, and paid/reimbursed amounts. It should be noted, however, that NSOC is a supplement to the National Health and Aging Trends Study (NHATS)²³ that collects information on caregivers to NHATS participants (i.e., Medicare beneficiaries). Medicare claims are therefore linked to NHATS data, which can then be linked to NSOC data. Linked PSID-Medicare data and linked MEPS-Medicare data cover all economic outcome domains. PSID includes five economic measures—cost of informal caregiving, absenteeism, time costs of home production and leisure, out-of-pocket health care costs, and paid/reimbursed

amounts—related to both patient and caregiver perspectives. MEPS includes four economic measures—cost of informal caregiving, absenteeism, insurance premiums, and out-of-pocket health care costs—related to all stakeholder perspectives.

Exhibit 7. Economic Outcome Measures in Federally Funded Data Sources Linked or Linkable to Medicare Claims

Outcome Measure	Number of Data Sources	% of Data Sources ¹
Domain: Direct Medical Costs		
Paid/reimbursed amount	10	56%
Out-of-pocket health care costs and/or medical expenses	11	61%
Insurance premiums	8	44%
Domain: Direct Non-Medical Costs		
Transportation and/or travel costs associated with seeking medical care	2	11%
Vehicle modification expenses	0	0%
Paid professional care (child care expenses, senior care expenses, and/or housekeeping expenses)	1	6%
Relocation/moving costs	0	0%
Specialized clothing/laundry costs	0	0%
Value of time spent in seeking health care	2	11%
Value of informal caregiving	3	17%
Special food	2	11%
Home modifications	2	11%
Domain: Indirect Costs		
Value of absenteeism	5	28%
Value of presenteeism	2	11%
Value of time spent in home production and leisure	4	22%
Lost wages from inability to work	4	22%

Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

¹ Totals do not add up to 100 percent due to non-mutual exclusivity of categories.

Features of the Resulting Medicare FFS Claims-Based Linked Data

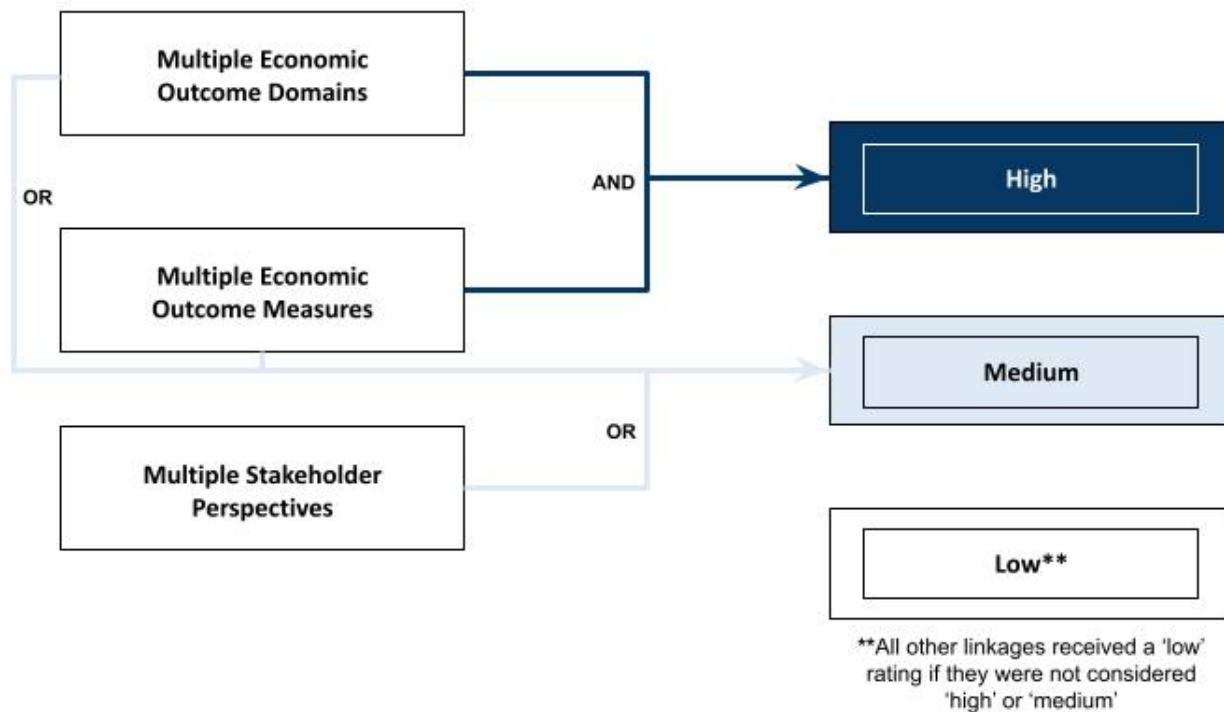
Since linked data create a higher risk of re-disclosure, additional access requirements apply for linked data since permissions are often required not just by CMS (for the Medicare FFS data), but also by the owner of the other data source. For example, to access the linked PSID-Medicare claims files, researchers must have both CMS and PSID data use agreements and additional restricted access provisions (a virtual data environment, managed by PSID). As another example, Medicare data linked to selected federal health data sources, the National Health and Nutrition Examination Survey²⁴ or the National Health Interview Survey²⁵, are available to researchers through the National Center for Health Statistics (NCHS) Research Data Center, conditional on application, approval, and specific terms.²⁶ Some key informants noted that processing times are a significant barrier to proposing feasible research plans for grants and to conducting timely analyses. In some cases, linkages also require an access fee. For example, informants noted that costs to use CMS’s VRDC—a virtual research

environment where claims linkages can be stored and analyzed—may be prohibitive for some researchers, depending on the research questions and volume of data being used. In particular, high access costs may dissuade early career investigators from requesting for linked data, raising questions of equity.

The time periods covered by the linkages significantly vary. Some sources contain a decade or more of linked data, while others are linked only for one year or a subset of years. Linkages with the longest time frame of available data include the Medicare Current Beneficiary Survey (29 years),²⁷ the Health and Retirement Study (HRS) (30 years),²⁸ and the PSID, the National Health Interview Survey (NHIS), and the National Health and Nutrition Examination Survey (NHANES) (each 19 years).

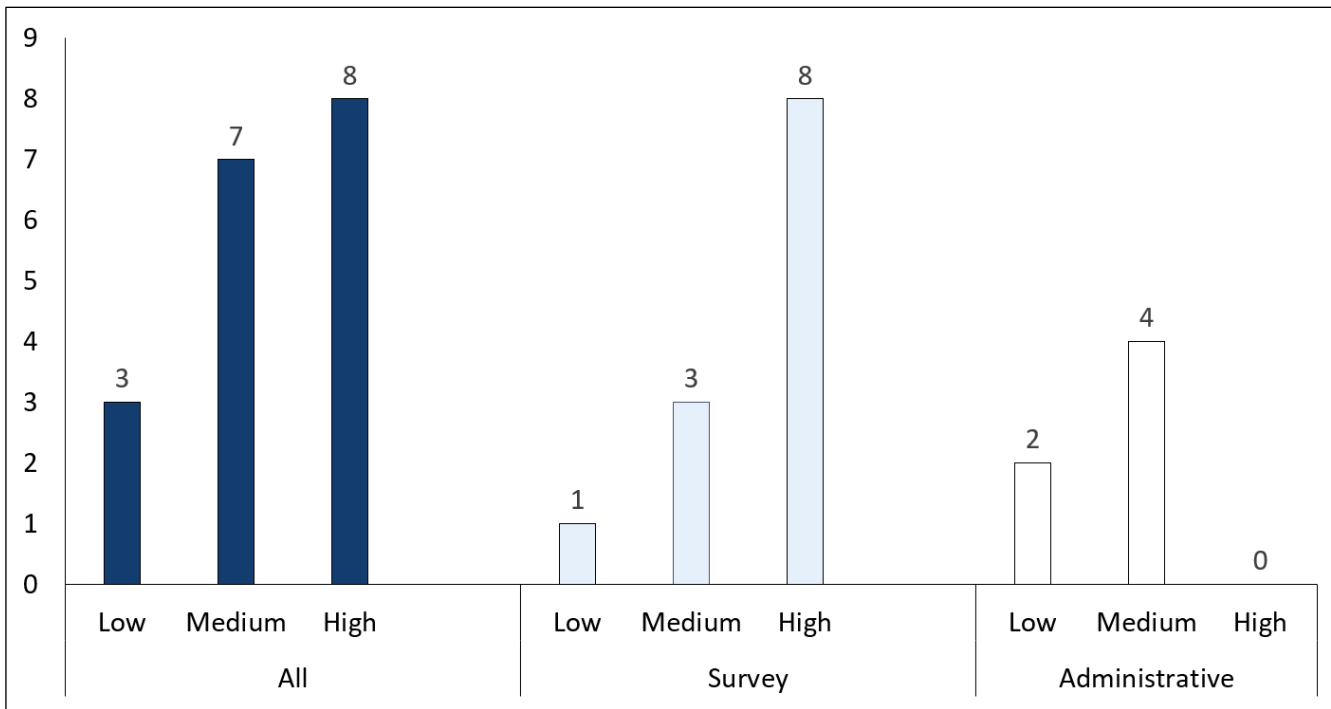
We rated the potential of each resulting linkage for PCOR economic evaluations as high, medium, or low, based on the extent of economic information available in each. [Exhibit 8](#) shows the rating rubric that was applied to the inventory. As shown, a “high” rating was applied when multiple economic outcome domains and multiple perspectives are available in a linkage; “medium” was used when there is at least one element of multiplicity (e.g., multiple economic outcome domains or multiple perspectives or multiple outcome measures within a given outcome domain); and we assigned a “low” rating for all other resulting linkages.

Exhibit 8. Rubric to Evaluate the Potential of Claims-Based Linkages for PCOR Economic Evaluations



Based on this scale, as shown in [Exhibit 9](#), eight data sources received a “high” rating. Three linkages received a “low” rating. No linkage of Medicare claims with other administrative data sources received a high rating.

Exhibit 9. Potential of Resulting Linkages for PCOR Economic Evaluations, by Source Type



Note: The findings in this exhibit are based on the 18 federal and federally funded data sources identified using the search process described earlier. Although a comprehensive scan was conducted, the exhibit may not represent all datasets that are linked or linkable to Medicare claims.

Strengths and Limitations of the Current Medicare-Based Linkages Infrastructure

[Exhibit 10](#) lists the relative strengths and weaknesses of survey versus administrative data that can be linked to Medicare claims. As seen, linking survey data to Medicare claims has several advantages for conducting PCOR economic studies versus using one or the other data type alone. Most surveys on their own have limited data on individuals' health and economic outcomes due to self-reported measures. Therefore, a key advantage of linking survey with claims data is to better attribute health care treatments or services received by patients to economic outcomes (via diagnosis and procedure codes). When Medicare claims are linked specifically with longitudinal survey data (e.g., PSID, National Long Term Care Survey,²⁹ NSOC, NHATS, and the Medicare Current Beneficiary Survey [MCBS]), outcomes and utilization can be analyzed over time using robust statistical and causal inference methods. Finally, as noted earlier, surveys capture a wider range of economic outcomes, including indirect costs and direct non-medical care costs, as well as SDOH and equity measures that are not present in claims or other administrative data sources. Therefore, linked survey and Medicare claims data can provide a more comprehensive picture not only of health outcomes but also of economic outcomes in the context of individuals' social conditions.

Exhibit 10. Strengths and Weakness of Resulting Linkages

Linkage Type	Strengths	Weaknesses
Survey Data-Based Linkages with Medicare Claims	<ul style="list-style-type: none"> ■ Enhanced attribution of outcomes to services and procedures ■ Wider range of economic outcome measures and SDOH/equity measures 	<ul style="list-style-type: none"> ■ Smaller sample sizes ■ Data lags ■ Shifting survey designs limiting the potential for longitudinal or trend analyses
Administrative Data-Based Linkages with Medicare Claims	<ul style="list-style-type: none"> ■ More specific disease information ■ Can study dual eligibles ■ Large sample sizes ■ Granular geographic information ■ Longitudinal analyses 	<ul style="list-style-type: none"> ■ Limited SDOH/equity information ■ Limited time frames for some of the linked data

There are also limitations with linked survey and Medicare claims data. Some surveys are not limited to Medicare populations (e.g., PSID), and linkages to Medicare data result in reduced sample sizes, which may limit analyses of subgroups and rare health events. Additionally, most survey data lag by two or more years before being made available to the public, whereas Medicare claims data are more current. Lags can also appear in linked survey and Medicare data due to the time needed to conduct and finalize linked data files. Researchers seeking to understand the most timely economic outcomes will be limited to the most currently available linked data, which will even under the most optimal conditions lag at least one to two years. Finally, survey sampling designs, questionnaires, and other features may change over time, limiting the ability to perform longitudinal or trend analysis even when linked datasets may span multiple years.

Linkages of other administrative data sources to Medicare claims also have several strengths for conducting PCOR economic studies. Existing linkages to disease registries (i.e., Surveillance, Epidemiology, and End Results [SEER]³⁰ and the United States Renal Data System³¹) allow researchers to gain more specificity about the nature of diseases (e.g., tumor characteristics) that are not captured via administrative codes, while linkages to Medicaid data allow researchers to conduct analyses related to the dual-eligible population. Most administrative linkages contain information on millions of individuals across the U.S. and include detailed geographic information (e.g., ZIP code or census tract). If individuals remain eligible for or enrolled in Medicare throughout the linked time period, robust longitudinal analyses may be conducted.

However, a primary weakness of administrative linkages is the limited information on SDOH and equity measures and lack of information on direct non-medical or indirect costs. Researchers seeking fuller pictures of outcomes or associations in these areas may not find utility in using linked, exclusively administrative datasets. Additionally, although linked administrative data sources often include large sample sizes and cover a national area, time frames may be limited. For example, data from the Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS) are available for only six years and from more than a decade ago (i.e., 2006-2012). SEER-Medicare and linked Medicaid Research Identifiable Files (RIFS) cover the longest and one of the most recent time frames (i.e., 1999-2019 for both).

Discussion

Linkages with Medicare claims are of particular interest for improving the PCOR economic outcomes landscape due to the large number of Medicare beneficiaries, the influential role of Medicare in the larger U.S. health care system, and the federal structure of the Medicare program. Following the scope and definitions described above,

this paper found that there are 18 separate data sources that are currently linked or linkable to Medicare claims data and can potentially be used for PCOR economic evaluations. At the same time, there are also numerous gaps in the linked data landscape that may be addressed in the future.

Considerations

Gaps in the Medicare linked data landscape are both missed opportunities for improved PCOR and areas that could be targeted for potential PCOR data infrastructure improvements in the future. Arguably the most important gap that we found is that none of the identified administrative data sources contain information on direct non-medical or indirect costs. This may result from the fact that such costs are often obtained from survey questions, though administrative measures may provide greater detail and accuracy than survey measures, and there are examples of measuring these outcomes in non-linked administrative data.^{32,33} As a result of this limitation, we also found that no linkage of Medicare claims with other administrative data sources received a “high” rating in our classification system.

We also found that none of the linked data sources covered certain economic outcomes. None of the linked sources provided information on relocation or moving costs, vehicle modifications, or specialized clothing or laundry costs, all of which fall under the domain of direct non-medical costs. Beyond Medicare linkages, these outcomes are, admittedly, reported in fewer PCOR studies overall, but they could still be highly relevant for certain patients.

The NSOC, MEPS, and PSID are the most promising for PCOR economic evaluations when linked to Medicare data, as they provide the largest number of measures and multiple perspectives for more comprehensive PCOR research. For common health problems or medical treatments, and interest in the included domains, these sources are very useful. However, they also have weaknesses. The MEPS is a nationally representative sample of households, capturing the civilian non-institutionalized population of all ages. Out of roughly 35,000 people per year, fewer than 5,000 MEPS participants are ages 65 and older, so the linked Medicare population is relatively small. The overlapping panel design of the MEPS also presents opportunities and limitations. While it provides frequent replenishment of new sample members, longitudinal follow-up of survey measures is limited to two years. The NSOC is nationally representative, but includes fewer than 10,000 participants per year, making it difficult to study rarer health outcomes or treatments. PSID is also nationally representative and is slightly larger overall in a single survey year but has some other tradeoffs. For example, PSID has fewer health-related survey questions than NSOC since health is not the primary purpose of the study. The PSID design is also concentrated on longitudinal follow-up, repeated observations of the same cohort over time. PSID and the HRS contain more detailed employment, income, and asset information than other data sources, making these uniquely useful for certain types of patient-centered studies.

The NSOC, MEPS, and PSID seem the most promising for PCOR economic evaluations when linked to Medicare data, based on the extent of economic measures that are available in them. Compared to the other identified data sources, these three provide the greatest number of economic measures and represent multiple stakeholder perspectives.

Only direct medical costs are available in the identified administrative data sources that are linked or linkable to Medicare claims. None of them contain information on direct non-medical or indirect costs.

An important future consideration for improving the landscape of PCOR economic outcomes research is the access to data. As with non-linked Medicare claims, no linkages of individual-level data (whether based on auxiliary survey or administrative data sources) are publicly available. All require research applications and/or data use agreements with the distributing entity. While protections for privacy, Health Insurance Portability and Accountability Act (HIPAA), and information security are necessary, this presents a significant barrier to research, and new innovations may be needed to improve access to linked data and to encourage PCOR research. Emerging solutions, such as synthetic data (e.g., Synthetic Healthcare Database for Research [SyH-DR])³⁴ or protected data enclaves (e.g., MedRIC)¹⁹, may provide such platforms and should be explored in future evaluations.

These considerations present opportunities for closing the identified gaps. Any efforts to reduce barriers to creating new linkages could greatly enhance the landscape of PCOR economic research, benefiting patients, payers, Medicare policymakers, researchers, and other stakeholders. This would be most fruitful if concentrated on areas where the gaps in measured outcomes are larger. Still, even efforts that expand existing areas can benefit patients and others greatly if they lead to the study of samples or populations not previously covered. For example, the NSOC, PSID, and HRS are powerful and representative, but are sampled from a small number of Medicare beneficiaries. While not sources of economic information themselves, informants also emphasized the importance of linking claims to supplemental sources that support better understanding of the drivers of costs. For example, informants mentioned enriched clinical information (e.g., from electronic health records), information on patients' social and economic context, and plan cost-sharing rules and network information as priorities for future linkages. Informants also mentioned non-federal linkages that are outside the scope of this paper but still potentially relevant for PCOR economic studies, including to proprietary or commercial data sources (e.g., a major health insurer) or to linkages maintained by university-led research teams (e.g., the Southern Community Cohort Study).

It should be noted that we also excluded geographic linkages, where information is added at aggregate levels, such as combining ZIP code or county-level information from external sources (such as the American Community Survey) with ZIP code data on Medicare claims. Such merges are common and important in health services research but typically do not add individual, patient-level detail on economic outcomes, so they were outside the scope of this paper. However, they may provide area-level economic information that could be useful for some PCOR analyses, and some of the equity and SDOH factors listed in our findings could be enhanced by including these.

Finally, while this paper focused on Medicare FFS claims, some amount of information about Medicare Advantage beneficiaries can be derived from some of the data linkages noted in this paper. For example, HRS provides survey and Master Beneficiary Summary File (MBSF) information for all survey respondents who consent to the Medicare claims linkage. Similarly, MBSF information is available for NHATS participants within the NHATS-CMS linked data, which can be used to identify Medicare Advantage beneficiaries. Additionally, Medicare Advantage encounter data have also been linked to other data sources (e.g., NHIS, NHANES, the National Nursing Home Survey [NNHS],³⁵ and the National Hospital Care Survey [NCHS]),³⁶ which provides researchers with opportunities to conduct economic analyses on Medicare Advantage beneficiaries using outcomes available in these survey data sources. NHIS and NHANES contain measures of indirect costs, enabling assessments of how health care utilization may impact these economic outcomes, and particularly for minority populations. Notably, NHIS and NNHS also contain information on out-of-pocket costs, and NHIS also contains some information on insurance premiums, which may help researchers better understand direct costs to

Medicare Advantage beneficiaries and how they may compare against costs incurred by their fee-for-service counterparts using NHIS- or NNHS-Medicare claims linked data.

Study Limitations

Our review is subject to some limitations. First, we relied on the economic outcomes identified in the previous report published by ASPE¹², which we believe is comprehensive but may still exclude certain perspectives or less common economic measures. Second, the scope of this work centered on federally funded linkages to Medicare, given the large size and relevance of this population. However, strictly private and state-specific sources may capture aspects of the Medicare population in other ways and would be missed by this work. We also cannot rule out the possibility of “one-off” projects in which individual researchers may have been able to combine their own primary data collection or clinical records with Medicare claims, leading to linkages that advance the PCOR landscape. Third, even among federal data, our search may have missed certain data sources, or specific economic or PCOR measures within the data sources. Finally, while discussions with subject matter experts were an invaluable resource, like the research team, they represent a set of perspectives from professional and personal experiences, which may not necessarily be representative of all perspectives that are important for PCOR.

Conclusion

Data on patient-centered economic outcomes related to health and health care are important for patients and their families. Complex health care decisions frequently present challenges for patients, families, providers, and other stakeholders; many such decisions may also have significant economic consequences. Improved research on patient-centered outcomes, including economic factors, associated with a range of health care topics and treatments, can provide invaluable information to help all types of stakeholders make better decisions, improving lives and health outcomes. Yet, to date, good information and widespread PCOR is often limited by available data sources and a fragmented health care system.

Medicare data represent an invaluable, large-scale resource for conducting PCOR studies of economic outcomes for nearly all older adults in the U.S., and many younger persons with disabilities and end-stage renal disease (ESRD). While Medicare FFS claims data have several advantages, their scope is limited by a small number of economics outcomes. Consequently, data linkages hold great potential for improving the federal data landscape on patient-centered research and economic outcomes. Barriers to access, high costs, and generalizability of some linkages inhibit wider research. However, efforts to continue to improve access and foster new data linkages may encourage methodological innovations by PCOR researchers, and help to produce important applied research for patients, caregivers, and policymakers. In particular, given the growing number of beneficiaries in the Medicare Advantage program, there is an urgent need for data that can be used to measure economic outcomes for this population. At present, all Part C data that are federally housed are encounter data, which are useful for measuring health care utilization but not costs. As Part C enrollment steadily grows, it will be increasingly worthwhile to prioritize identifying innovative ways in which encounter data can be used for patient-centered economic evaluations, as well as identifying how other data sources can be integrated with encounter data to study the economic impacts of health care treatments and services received by Medicare Advantage beneficiaries.

Appendix A: Economic Outcomes Relevant to PCOR Analysis of Health Care Interventions

Economic Outcome Category	Economic Outcome Measure(s)	Economic Outcome Description	Economic Outcome Example(s)	Patient/Family Perspective	Caregiver Perspective	Employer Perspective	Payer/Insurer Perspective
Direct Medical Costs	Paid/reimbursed amounts	Amount that insurer pays a provider for any type of covered medical care	Paid for physician, inpatient, outpatient, pharmacy/drug, lab			X	X
	Out-of-pocket (OOP) health care costs	Amount that individuals are responsible for after insurance	Copayments, coinsurance, deductibles, etc. associated with any covered service, including (1) costs associated with adverse events due to medical treatment, and (2) alternative medical treatments (e.g., acupuncture, massage, psychotherapy, personal trainers, and health coaches)	X			
	Other OOP medical expenses	Health-related expense not covered by an insurance plan	Over-the-counter drugs, other services not covered by insurance including (1) costs associated with adverse events due to medical treatment, (2) alternative medical treatments (e.g., acupuncture, massage, psychotherapy, personal trainers, and health coaches), or (3) genetic testing	X			
	Insurance premiums	Premium cost to purchase any type of health insurance	Health, dental, vision, drug, long-term care insurance, etc.	X		X	
Direct Non-Medical Costs	Transportation costs	Transportation costs associated with seeking medical care	Bus, taxi, car, mileage, gas, parking, etc.	X	X		
	Travel costs (non-transportation)	Other non-transportation travel costs associated with seeking medical care	Hotel, food, etc.	X	X		
	Vehicle modification costs	Costs associated with accessibility changes to personal vehicle	Wheelchair lift for a car	X	X		

Economic Outcome Category	Economic Outcome Measure(s)	Economic Outcome Description	Economic Outcome Example(s)	Patient/ Family Perspective	Caregiver Perspective	Employer Perspective	Payer/ Insurer Perspective
	Child care and other child-related expenses	Paid care when patient cannot do this due to medical needs	Babysitter, nanny - in or out of home, paid tutoring, or homeschooling	X	X		
	Housekeeping costs	Paid care when patient cannot do this due to medical needs	Housecleaner	X	X		
	Elder care costs	Paid care when patient cannot do this due to medical needs	Elder care - in or out of home	X	X		
	Special food costs	Costs associated with dietary modifications for medical treatment or condition	Prep for colonoscopy	X	X		
	Home modification costs	Costs associated with accessibility changes at home	Bath, shower, stairs, ramps	X	X		
	Relocation/moving costs	Costs incurred from a long-term, permanent move associated with treatment or condition	Cost of rent or mortgage associated with moving into a more accessible home, educational costs incurred by patient's family members (e.g., children) because of relocation	X	X		
	Clothing/laundry costs	Costs associated with specific modifications needed for medical treatment or condition	Hospital gown for home use	X	X		
	Time costs: Health care seeking	Any time-related costs associated with health service, medical treatment, or condition	Time in doctor's office, waiting, receiving care, etc.	X	X		
	Time costs: informal caregiving	Time spent caring for patient needs as a result of their medical treatment or condition	Value of assistance (time) from an unpaid relative or friend		X		

Economic Outcome Category	Economic Outcome Measure(s)	Economic Outcome Description	Economic Outcome Example(s)	Patient/ Family Perspective	Caregiver Perspective	Employer Perspective	Payer/ Insurer Perspective
Indirect Costs	Lost wages from absenteeism	Work loss (days/hours) associated with medical treatment or condition	Missed work	X	X	X	
	Lost wages from reduced productivity (presenteeism)	Reduced productivity on the job, associated with medical treatment or condition	Less productive on the job	X	X	X	
	Time Costs: Home Production and Leisure	Reduced home productivity and leisure associated with medical treatment or condition	Activity limitation days	X	X		
	Lost wages and benefits from losing work	Voluntary unemployment, involuntary unemployment, early retirement, or exit from labor force due to disability associated with medical treatment or condition	Labor force participation; retirement; permanent disability affecting work or career and associated earnings and benefits	X	X		

Appendix B: Key Informant Discussion Questions

Below are listed the questions that were used in the key informant interviews with researchers/data stewards and Medicare policymakers. In some cases, questions were slightly modified for the Medicare policymakers group, and indicated as such in italics.

1. I want to start by having you do a brief introduction of yourself. Please introduce yourself and tell us about your role within {the organization}.
 - a. Probe: Please briefly describe your background and familiarity with Medicare claims data, any other survey or administrative data that can be linked to Medicare, PCOR, comparative effectiveness research, or economic evaluations?

Existing Linkages Between Medicare Claims and Other Data Useful for PCOR Economic Evaluations

2. Does our inventory look comprehensive to you based on your knowledge of available survey and administrative data sources that have economic outcomes and that can be linked to Medicare claims? Are there any obvious ones that are missing?
 - a. Probe: Is a linked version between <this data source> and Medicare claims available to external researchers in a standardized format (“off the shelf”), or do researchers have to perform the linkage “from scratch” using common identifiers?
 - b. Probe: Are there any characteristics of <this data source> that makes it especially appealing for PCOR economic evaluations? What particular aspects?
3. Are there any data sources from our list that, based on your knowledge or experience, stand out as being the most useful for assessing Medicare beneficiaries’ economic outcomes? If so, why?

{Question version for Medicare Policymaker}: Are there any data sources from our list that, based on your knowledge or experience, stand out as being the most useful for assessing Medicare beneficiaries’ economic outcomes or for the work of federal policymakers? If so, why?

- a. Probe: Are there particular advantages of <this data source> in terms of its accuracy, timeliness, or other characteristics, over others?

Limitations or Gaps in the Existing Medicare-Based Linked Data Infrastructure for PCOR Economic Evaluations

4. Thinking specifically around linked Medicare data at large, what are some of the most basic or common linkage-related issues or hindrances that may inhibit health services research for the Medicare population, and that you think the Federal government should try to address in the near term?
 - a. Probe: For example, what threats exist which may actually lead to a decrease in research output on PCOR economic evaluations?
5. Looking, say, 5+ years into the future, will there be changes in how researchers use or conduct research from Medicare claims-based linked data? In what ways?

{Question version for Medicare Policymaker}: Looking, say, 5+ years into the future, will there be changes in how federal policymakers use or conduct research from Medicare claims-based linked data? In what ways?

- a. Probe: For example, do you anticipate that researchers would primarily work within virtual environments (e.g., VRDC, MedRIC, Census RDCs) so that there is reduced handling of Protected Health Information (PHI)? What about accessing data via a Tokenization platform like that provided by Datavant?
 - b. Probe: Are there any data access or use considerations that make you more optimistic about being able to increase the output of meaningful PCOR economic evaluations for the Medicare population?
 - c. Probe: What about any data access or use considerations that make you less optimistic about being able to increase the output of meaningful PCOR economic evaluations for the Medicare population?
6. Generally speaking, can you think of any new priority areas or considerations for enhancing the availability, access, and value of existing Medicare claims-based data linkages for PCOR-based economic evaluations, particularly for external researchers?

{Question version for Medicare Policymaker}: *Let's suppose that a national policymaking group were to make some recommendations around new priority areas or considerations for enhancing the availability, access, and value of existing Medicare claims-based data linkages for PCOR-based economic evaluations. What kinds of things might the group identify?*

- a. Probe: What are some reasonable priorities that the Federal government could consider in the near term to promote availability and access?
- b. Probe: In your assessment, what are some of the data infrastructure improvement efforts that the external research community in particular would most value and appreciate?
- c. Probe: What data improvement efforts in particular can enhance the potential for equity-focused PCOR-based economic evaluations?

Forthcoming Linkages Between Medicare Claims and Other Data Useful for PCOR Economic Evaluations

7. Are you aware of any emerging Medicare claims-based data linkages (with respect to survey or administrative data) that may become available to external researchers in the future, and that could be used for PCOR-based economic evaluations for the Medicare population?
- a. Probe: What is the status of <this effort>, if you know? What is the projected time frame for availability of <this linkage>?
 - b. Probe: Are you able to speak generally to the degree to which economic outcomes would be available via this linkage? For instance, what types or categories of economic outcomes do you know are available in the auxiliary <data source>?
 - c. Probe: Do you know if the plan is to make a pre-linked version available to researchers, or just to enable the linkage between <this data source> and Medicare claims?
 - d. Probe: Will there be any access limitations faced only by outside researchers versus their counterparts in the Federal government?

8. Can you think of any data sources or data types that are currently not linked or linkable to Medicare claims, but would be valuable to the research community going forward for conducting PCOR economic analyses for the Medicare population?

{Question version for Medicare Policymaker}: *Can you think of any data sources or data types that are currently not linked or linkable to Medicare claims, but would be valuable to the research community going forward for policy-relevant PCOR-based economic research for the Medicare population?*

- a. Probe: Are there any data sources that you have or are aware of (e.g., primary survey data, clinical trials data) that you would like to see linked or be linkable to Medicare claims but that currently are not?
- b. Probe: Are you aware of any current efforts (within or outside the Federal government) that are aimed at closing this gap in data availability?
- c. Probe: What steps can the Federal government take to enable their availability and access, particularly for outside researchers?

Research Suggestions Regarding Federal Data Linkages for Economic Variables

9. What areas of economic burdens and impacts relevant to the Medicare population need more study, but are not possible with today's linked data sources? Importantly, if you have done or are doing any economic analyses or evaluations that are based on outcomes that we haven't considered, we would love to hear those.
 - a. Probe: What kinds of data sources or linkages could facilitate the assessment of such outcomes?
 - b. Probe: Are you aware of any discussions or plans (either within or outside of the Federal government) to make such data sources and linkages available to external researchers?
10. And finally, I am interested to hear any suggestions you might have for other possible white papers that describe the availability of Medicare, Medicaid and other federal data linkages for PCOR-based economic evaluations. For example, are there specific populations or diseases that ASPE should focus on, for future research papers?
 - a. Probe: Do you have any suggestions for white papers that might describe the availability of federal data for equity-focused PCOR economic evaluations?

Appendix C: Federally funded Data Sources Currently LINKED OR Linkable to Medicare FFS Claims

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
Medicare Part D claims (PDE) CMS	https://resdac.org/cms-data/files/pde	Administrative	48 million	Direct Medical Costs	<u>Direct Medical Costs:</u> Insurance premium*; OOP health care costs, Paid/reimbursed amount <i>*Premium information is available in the Plan Characteristics file</i>	2006-2020	Via research application to ResDac	Medium	Multiple perspectives represented with the linkage
Medicare Current Beneficiary Survey (MCBS) CMS	https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS	Survey	11,548	Direct Medical Costs, Direct Non-Medical Costs	<u>Direct Medical Costs:</u> Insurance Premium, OOP health care costs, Paid/reimbursed amount <u>Direct Non-Medical Costs:</u> Transportation, Time costs: health care seeking	1991-2020	Via research application to ResDAC.	High	Multiple economic outcome domains and perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
Medicaid claims (Research Identifiable Files) (RIF) CMS	<p>MAX: https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAXGeneralInformation</p> <p>TAF: https://www.medicaid.gov/medicaid/data-systems/macbis/medicaid-chip-research-files/transformed-medicaid-statistical-information-system-t-msis-analytic-files-taf/index.html</p>	Administrative	Varies by file type; 73+ million Medicaid; 6.7 million CHIP	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount, OOP health care costs	1999-2019	Via research application to ResDAC	Medium	Multiple economic measures and perspectives represented with the linkage
Medicare-Medicaid Linked Enrollee Analytic Data Source (MMLEADS) CMS	https://resdac.org/cms-data/files/mmleads-1	Administrative	58+ million	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount	2006-2012	Via research application to ResDAC	Low	Only one economic domain and one perspective represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
National Health Interview Survey (NHIS) CDC/NCHS	https://www.cdc.gov/nchs/data-linkage/medicare.htm	Survey	30,000	Direct Medical Costs, Indirect Costs	<u>Direct Medical Costs:</u> Insurance Premium, OOP health care costs <u>Indirect Costs:</u> Absenteeism, Time costs: home production and leisure, Inability to work, Productivity	1994-2018 NHIS data has been linked to 1994-2018 Medicare enrollment data and 1994-2013 and 2016-2018 FFS claims.	Via research application to NCHS Research Data Center (RDC)	High	Multiple economic outcome domains and perspectives represented with the linkage
National Health and Nutrition Examination Survey (NHANES) CDC/NCHS	https://www.cdc.gov/nchs/data-linkage/medicare.htm	Survey	9,254	Direct Non-Medical Costs, Indirect Costs	<u>Direct Non-Medical Costs:</u> Special food <u>Indirect Costs:</u> Time costs: home production and leisure	1999-2018 Continuous NHANES and Third NHANES (NHANES III) data has been linked to 1999-2018 Medicare enrollment data and 1999-2013 and 2016-2018 FFS claims	Via research application to NCHS RDC	High	Multiple economic outcome domains and perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
The Second Longitudinal Study of Aging (LSOA II) CDC/NCHS	https://www.cdc.gov/nchs/data-linkage/medicare.htm	Survey	7,527	Indirect Costs	<u>Indirect Costs:</u> Absenteeism, Time costs: home production and leisure, Inability to work, Productivity	The 1994 LSOA II survey data are linked to 1991 – 2013 Medicare FFS claims.	Via request of the CD-ROM or diskettes; e-mail nchsquery@cdc.gov or telephone (301) 458-INFO.	High	Multiple economic outcome domains and perspectives represented with the linkage
National Hospital Care Survey (NHCS) CDC/NCHS	https://www.cdc.gov/nchs/data-linkage/nhcs-linkage.htm	Survey	500 hospitals	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount	2014, 2016	Via research application to NCHS RDC	Low	Only one economic domain and one perspective represented with the linkage
The National Nursing Home Survey (NNHS) CDC/NCHS	https://www.cdc.gov/nchs/data-linkage/medicare.htm	Survey	1 million+	Direct Medical Costs	<u>Direct Medical Costs:</u> OOP health care costs	2004 NNHS has been linked to 1999-2018 Medicare enrollment data and 1999-2013 and 2016-2018 FFS claims	Via research application to NCHS RDC	Medium	Multiple economic measures and perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
United States Renal Data System (USRDS) NIH	https://usrds.org/	Administrative	1 million+	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount, OOP health care costs	2011-2019	Via a Research Proposal and a USRDS Merged Dataset Agreement for Release of Data	Medium	Multiple economic measures and perspectives represented with the linkage
National Health and Aging Trends Study (NHATS) NIH	https://www.medric.info/partners-pages/nhats	Survey	16,283	Direct Medical Costs, Indirect Costs	<u>Direct Medical Costs:</u> Insurance Premium <u>Indirect Costs:</u> Inability to work	Longitudinal linked data available for participants in the 2011 cohort and new participants in the 2015 cohort	NHATS-CMS linked data files are available through the Health and Aging Data (HaAD) Enclave. Access to linked data requires a DUA with NHATS and data protection plan.	High	Multiple economic outcome domains and perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
National Study of Caregiving (NSOC) NIH	https://nhats.org/researcher/nsoc	Survey	2,361	Direct Non-Medical Costs, Indirect Costs	<u>Direct Non-Medical Costs:</u> Time costs: Informal caregiving, Transportation, Special food, Time costs: healthcare seeking, home modifications, housekeeping <u>Indirect costs:</u> Absenteeism, Inability to work	Longitudinal linked data available for participants in the 2011 NHATS cohort and new participants in the 2015 NHATS cohort.	NHATS-CMS linked data files are available through the Health and Aging Data (HaAD) Enclave. Access to linked data requires a DUA with NHATS and data protection plan. The NSOC is considered NHATS Restricted Data and can be requested along with the CMS-linked data.	High	Multiple economic outcome domains and perspectives represented with the linkage
National Long Term Care Survey (NLTCs) NIH	https://www.icpsr.umich.edu/web/NACDA/studies/9681/summary	Survey	35,789	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount, OOP health care costs, Insurance Premium	1984-2004	Via research application to the National Archive of Computerized Data on Aging. Access to linked data requires a DUA	Medium	Multiple economic measures and perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
SEER-Medicare NIH*	https://healthcaredelivery.cancer.gov/seermedicare/	Administrative	6 million+	Direct Medical Costs	<u>Direct Medical Costs:</u> Paid/reimbursed amount	1999-2019	Via research proposal to IMS (NCI's information technology contractor).	Low	Only one economic domain and one perspective represented with the linkage
Health and Retirement Study (HRS) NIH	https://resdac.org/cms-data/files/hrs-medicare	Survey	27,895	Direct Medical Costs	<u>Direct Medical Costs:</u> Insurance Premium, Paid/reimbursed amount, OOP health care costs	1991 - 2021	First step is a research application to HRS. Additional steps: - If research project is funded by NIA, researcher should apply to MedRIC to access the data through the MedRIC Enclave. - If research project is not funded by the NIA, researcher will apply for a CMS DUA through ResDAC before requesting the data from MedRIC	Medium	Multiple perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
Panel Study of Income Dynamics (PSID) <i>University of Michigan</i>	https://psidonline.isr.umich.edu/	Survey	26,000	Direct Medical Costs, Direct Non-Medical Costs, Indirect Costs	<u>Direct Medical Costs:</u> Insurance Premium, OOP health care costs <u>Direct Non-Medical Costs:</u> Informal Caregiving <u>Indirect Costs:</u> Absenteeism, Time costs: home production & leisure	1991-2010	Via a Data Use Agreement (DUA) from CMS and approval from PSID	High	Multiple economic outcome domains and perspectives represented with the linkage
Health Economics Resource Center Average Cost Datasets (HERC) <i>U.S. Department of Veterans Affairs</i>	https://www.herc.research.va.gov/include/page.asp?id=home	Administrative	Unknown	Direct Medical Costs	<u>Direct Medical Costs:</u> (Approximation of) Paid/reimbursed amount, OOP health care costs* *Estimates are average costs	NA	Outside researchers must collaborate with a VA researcher	Medium	Multiple perspectives represented with the linkage

Data Source Name (Abbreviation) Steward	URL for Linked Data	Source of Data	Sample Size	Available Economic Outcome Categories	Available Economic Outcome Measures, by Category	Available Time Frame for Linked Data	Outside Researchers' Access to Linked Data	Rating of Linked Data for PCOR Economic Evaluations	Rationale for Rating
Medical Expenditure Panel Survey (MEPS)** Agency for Healthcare Research and Quality	https://meps.ahrq.gov/mepsweb/	Survey	30,716 individuals, 12,756 families	Direct Medical Costs, Direct Non-Medical Costs, Indirect Costs	<u>Direct Medical Costs:</u> Insurance premium, OOP health care costs <u>Direct Non-Medical Costs:</u> Informal Caregiving <u>Indirect Costs:</u> Absenteeism	1996-2019 MEPS linked with 2014-2018 Medicare claims 1996-2014 MEPS linked with 1999-2013 Medicare claims	Via research application to the National Center for Health Statistics Research Data Center	High	Multiple economic outcome domains and perspectives represented with the linkage

* With the exception of the SEER-Medicare data which is already in linked form, the economic outcome measures field for all other data sources represents the specific outcomes that are available independently in the data source versus via the linkage.

** Although publicly available information about the MEPS-Medicare data linkage is not available, a technical expert panel (TEP) panel member at AHRQ confirmed with the research team that external researchers can request for MEPS-Medicare linked data through the NCHS RDC.

References

- ¹ CMS Releases Latest Enrollment Figures for Medicare, Medicaid, and Children’s Health Insurance Program (CHIP). Medicaid.gov. Accessed June 22, 2022.
<https://content.govdelivery.com/accounts/USCMSMEDICAID/bulletins/310d30e>
- ² Tarazi W, Welch WP, Nguyen N, et al. Medicare Beneficiary Enrollment Trends and Demographic Characteristics. Published online March 2, 2022:13.
- ³ Clemens J, Gottlieb JD. In the Shadow of a Giant: Medicare’s Influence on Private Physician Payments. *Journal of Political Economy*. 2017;125(1):1-39. doi:10.1086/689772
- ⁴ Foote SB, Virnig BA, Town RJ, Hartman L. The Impact of Medicare Coverage Policies on Health Care Utilization. *Health Serv Res*. 2008;43(4):1285-1301. doi:10.1111/j.1475-6773.2008.00836
- ⁵ Mues KE, Liede A, Liu J, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. *Clin Epidemiol*. 2017;9:267-277. doi:10.2147/CLEP.S105613
- ⁶ ResDAC. Strengths and Limitations of CMS Administrative Data in Research. Accessed July 25, 2022.
<https://resdac.org/articles/strengths-and-limitations-cms-administrative-data-research>
- ⁷ Hasche J, Ward C, Schluterman N. Diabetes Occurrence, Costs, and Access to Care among Medicare Beneficiaries Aged 65 Years and Over. :8.
- ⁸ Rein DB, Zhang P, Wirth KE, et al. The economic burden of major adult visual disorders in the United States. *Arch Ophthalmol*. 2006;124(12):1754-1760. doi:10.1001/archophth.124.12.1754
- ⁹ Data Dictionaries. Chronic Conditions Data Warehouse. Accessed September 12, 2022.
<https://www2.ccwdata.org/data-dictionaries>
- ¹⁰ Jo C. Cost-of-illness studies: concepts, scopes, and methods. *Clin Mol Hepatol*. 2014;20(4):327-337. doi:10.3350/cmh.2014.20.4.327 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4278062/>
- ¹¹ U.S. Department of Health and Human Services (HHS), Federal Coordinating Council for Comparative Effectiveness Research. Report to the President and Congress. June 30, 2009. Available at <https://osp.od.nih.gov/wp-content/uploads/FCCER-Report-to-the-President-and-Congress-2009.pdf> (accessed 26 May, 2021).
- ¹² Brown D, Srinivasan M, Arbulu L, Dullabh P, Curtis K, Zott C, and Smith SR. Federal Data for Conducting Patient-Centered Outcomes Research on Economic Outcomes. September 2021. ASPE Research Report.
- ¹³ 116th Congress. Further Consolidated Appropriations Act of 2020; Pub. L. No. 116-94 (2019), H.R. 1865.

¹⁴ Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost-Effectiveness in Health and Medicine. Second Edition, New to this Edition: Oxford University Press; 2016. Accessed August 30, 2021.

<https://global.oup.com/academic/product/cost-effectiveness-in-health-and-medicine-9780190492939?cc=us&lang=en&>

¹⁵ CMS. National Health Expenditure (NHE) Fact Sheet. Accessed June 22, 2022.

<https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet>

¹⁶ CMS Releases Final 2015 Medicare Advantage Encounter Data to Researchers | CMS. Published July 17, 2019. Accessed September 12, 2022. <https://www.cms.gov/newsroom/press-releases/cms-releases-final-2015-medicare-advantage-encounter-data-researchers>

¹⁷ Meyers DJ, Mor V, Rahman M, Trivedi AN. Growth In Medicare Advantage Greatest Among Black And Hispanic Enrollees. Health Affairs. 2021;40(6):945-950. doi:10.1377/hlthaff.2021.00118

¹⁸ The Shocking Truth About Medicare Advantage Plans. Medicare Plan Finder. Accessed August 1, 2022.

<https://www.medicareplanfinder.com/medicare-plans/medicare-advantage-plans/>

¹⁹ MedRIC. (2022). Medicare & Medicaid Resource Information Center (MedRIC). Medicare & Medicaid Resource Information Center. Accessed August 15, 2022. <https://www.medic.info/>

²⁰ Institute for Social Research. PSID Home. Accessed August 15, 2022. <https://psidonline.isr.umich.edu/>

²¹ Medical Expenditure Panel Survey Home. Accessed August 18, 2022. <https://meps.ahrq.gov/mepsweb/>

²² National Study of Caregiving (NSOC). Accessed August 15, 2022. <https://nhats.org/researcher/nsoc>

²³ National Health and Aging Trends Study (NHATS). Welcome Researchers. Accessed August 15, 2022.

<https://www.nhats.org/researcher>

²⁴ NHANES National Health and Nutrition Examination Survey. NHANES Homepage. Accessed August 15, 2022.

<https://www.cdc.gov/nchs/nhanes/index.htm>

²⁵ NHIS - National Health Interview Survey. Accessed August 15, 2022.

<https://www.cdc.gov/nchs/nhis/index.htm>

²⁶ NCHS RDC - Research Data Center Homepage. Accessed August 15, 2022.

<https://www.cdc.gov/rdc/index.htm>

²⁷ Medicare Current Beneficiary Survey (MCBS). Welcome to the MCBS! Accessed August 15, 2022.

<https://mcbs.norc.org/>

²⁸ The Health and Retirement Study (HRS). Accessed August 15, 2022. <https://hrs.isr.umich.edu/about>

²⁹ National Long Term Care Survey (NLTC). National Institute on Aging. Accessed August 15, 2022. <https://www.nia.nih.gov/research/resource/national-long-term-care-survey-nltcs>

³⁰ Seer-Medicare. Seer-Medicare Linked Data Resource. Accessed August 16, 2022. <https://healthcaredelivery.cancer.gov/seermedicare/>

³¹ The United States Renal Data System (USRDS) Accessed August 15, 2022. <https://www.usrds.org/>

³² Short ME, Goetzel RZ, Pei X, et al. How Accurate are Self-Reports? An Analysis of Self-Reported Healthcare Utilization and Absence When Compared to Administrative Data. *J Occup Environ Med.* 2009;51(7):786-796. doi:10.1097/JOM.0b013e3181a86671

³³ Mitchell RJ, Bates P. Measuring Health-Related Productivity Loss. *Popul Health Manag.* 2011;14(2):93-98. doi:10.1089/pop.2010.0014

³⁴ Synthetic Healthcare Database for Research (SyH-DR). Accessed August 15, 2022. Agency for Healthcare Research and Quality, Rockville, MD. <https://www.ahrq.gov/data/innovations/syh-dr.html>

³⁵ NNHS - National Nursing Home Survey. NNHS Homepage. Centers for Disease Control and Prevention. Accessed August 15, 2022. <https://www.cdc.gov/nchs/nnhs/index.htm>

³⁶ National Hospital Care Survey. About the NHCS. Centers for Disease Control and Prevention. Accessed August 15, 2022. <https://www.cdc.gov/nchs/nhcs/index.htm>