



Leibniz Institute
for Prevention Research and
Epidemiology – BIPS

Medicaid and Medicare

Dirk Enders, Tania Schink, Til Stürmer

DOI

10.1007/978-3-030-51455-6_19

Published in

Databases for Pharmacoepidemiological Research

Document version

Accepted manuscript

This is the author's final accepted version. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Online publication date

12 January 2021

Corresponding author

Tania Schink

Citation

Enders D, Schink T, Stürmer T. Medicaid and Medicare. In: Sturkenboom M, Schink T, editors. *Databases for Pharmacoepidemiological Research*. Cham: Springer International Publishing; 2021. p. 231-42.

Medicaid and Medicare

Dirk Enders, Tania Schink, Til Stürmer

1. Database description

1.1 Introduction

Medicaid is the largest health care program for persons with low income in the US and is jointly funded by the federal and individual state governments, while Medicare is solely funded by the federal government and provides health care for the vast majority of elderly persons. Both programs were established in 1965 and are overseen by the Centers of Medicare and Medicaid Services (CMS) of the United States Department of Health and Human Services. The CMS also maintain a database with administrative data from both health care programs, which is the main source for researchers working with Medicaid or Medicare data (Leonard et al. 2017). The administrative data of Medicaid has been used to answer pharmacoepidemiological research questions since the early 1980s (Hennessy et al. 2012a). Administrative data of Medicare has increasingly been used for this purpose since the implementation of prescription drug coverage in 2006 (Hanlon and Donohue 2010).

1.2 Database characteristics

Medicaid originally covered expenditures for low income pregnant women, families with children or elderly patients as well as chronically disabled patients in all federal states of the US. In 2014, the eligibility was extended to all low income patients in the course of the Patient Protection and Affordable Care Act, but extension is voluntary and varies by state. Medicare covers most patients aged 65 years and above as well as some disabled persons and patients with end-stage renal disease or amyotrophic lateral sclerosis. The Medicare program is divided into four parts. Part A is available to all Medicare enrollees and covers inpatient and hospice/nursing home care. Part B is available for an additional monthly fee and covers outpatient physician visits, services or products. Part C includes Medicare Advantage plans, which can be demanded by Part A or B enrollees and cover additional services, typically paid by an extra premium each month. However, the claims of Part C are generally not available for research through CMS. Part D was implemented in 2006 as part of the Medicare Modernization Act of 2003 and covers outpatient prescriptions. To be included in Part D, patients must enroll in stand-alone prescription drug plans or Medicare Advantage prescription drug plans, which are administered by private health insurances. As the eligibility criteria for Medicare and Medicaid overlap patients could be enrolled in both Medicare and Medicaid.

The trend of enrolled patients over time is depicted in Figure 1. There has been a steady increase in enrolled persons over the last 50 years in both health care programs with a steeper increase in enrollment for Medicaid than for Medicare (Part A or B). Table 1 shows the latest demographic characteristics. Medicaid covered 74.49 million people in 2012 and 58.48 million were enrolled in Medicare in 2017, corresponding to 23.7% and 18.0% of the

US population, respectively (U.S. Census Bureau Population Division 2018). The beneficiaries of Medicaid are not representative for the US population, as children, females as well as non-white persons are over-represented due to the selective eligibility criteria of the program. For the same reason, white persons, females and seniors are overrepresented in Medicare. Over 10.6 million patients are currently enrolled in both Medicare and Medicaid (Centers for Medicare & Medicaid Services 2020).

The raw administrative data of Medicaid is processed by the CMS and made available for research purposes via research identifiable files (RIFs). For Medicaid, the administrative data needs three to four years to become available and the RIFs currently cover data from 1999 up to 2013 for all federal states and up to 2015 for some of them. The administrative data of Medicare Part A, B, and D is updated more frequently than the Medicaid data and is now available after approximately one year. Currently, the RIFs of Medicare cover the years 1999 to 2017 and 2018 data is expected to be complete at the beginning of 2020 (Part D data is available from 2006 and already includes 2018 data)

1.3 Available data

Medicaid data is stored in five different data files, which are referred to as Medical Analytic eXtract (MAX) files by the CMS:

- The Personal Summary File contains demographic characteristics of the patients, e.g., date of birth, gender and race/ethnicity, date of death (without cause of death). It also contains information about the eligibility status (e.g., eligibility group, months of eligibility, dual eligibility to Medicare and Medicaid) and summary measures on use of the health system (e.g., number and duration of hospital stays, total number of prescribed drugs, payments).
- The Inpatient File contains all hospital stay records for enrollees using inpatient services. Included are hospital admission dates, type and begin/end of services, status at discharge and payments. Diagnoses are coded by the International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9-CM) and procedures by ICD-9-CM, Healthcare Common Procedure Coding System (HCPCS) level I-III or state-specific codes.
- The Other Therapy File includes claims from outpatient and inpatient physician visits (including physician specialty), as well as outpatient hospitals, clinics, home health care and hospices with respective dates. Diagnoses and procedures are coded in the same manner as in the Inpatient File. Outpatient laboratory and radiology records are also contained, but lab results are not reported.
- The Prescription Drug File covers all drug claims of Medicaid, which are coded according to the National Drug Code (NDC) system. Drugs identified by other codes

such as HCPCS or state-specific codes are contained in the Other Therapy file. Note that most drugs for patients which are dually eligible for Medicaid and Medicare are contained in the Medicare files. All drugs are contained with prescription date, prescription fill date, strength, quantity and duration of supply, and identification number of the prescribing physician. Non-reimbursable drugs administered at hospital and indications for drugs are not contained.

- The Long Term Care File includes claims from long-term care facilities, i.e., nursing homes, intermediate care facilities, and psychiatric facilities. The file includes admission dates to the facility, dates of services, diagnoses coded in ICD-9-CM, and discharge status of the patient.

The data of Medicare is available via numerous data files, which can be linked via unique subject identifiers:

- The Master Beneficiary Summary File contains demographic characteristics such as sex, race, region of residence, date and cause (only from 1999 to 2008) of death, monthly enrollment information for Plan C and D, and summary measures on costs and uses of services.
- The Standard Analytic Files (SAFs), also known as Medicare Claims Files, contain claims from institutional and non-institutional health care providers. Institutional data covers inpatient and outpatient data as well as claims from skilled nursing facilities, hospices, and home health agencies. Non-institutional claims cover durable medical equipment and data on physicians and free-standing facilities such as clinical laboratories. In general, diagnosis (coded in ICD-9-CM) and procedures (coded in ICD-9-CM or HCPCS) are included together with the date of service and the amount of reimbursement.
- The Medicare Provider and Analysis Review (MedPAR) Files contain inpatient hospital and/or skilled nursing facility final action claims with diagnoses and procedures (coded in ICD-9-CM) and the corresponding date of service and reimbursement amount. Contrary to the SAFs, each record represents a complete stay in a hospital or nursing facility and thus might contain multiple claims if they belong to the same stay.
- Information on prescription claims is available via the Part D Drug Event (PDE) File. This file contains one record per prescription and contains prescription date, NDC, days of supply and quantity dispensed. As in the Medicaid data, indications for drugs are missing. Further, medication administered at the hospital is covered by Plan A and is thus not contained in the PDE file. Information on prescription drug plans, pharmacies, drugs and prescribers is available in supplemental files.

Medicaid and Medicare data can be linked to each other but also to other data sources. For example, since laboratory results are absent, a study has been performed to link Medicare data of 10 eastern states data of a large national laboratory service (Hammill et al. 2015). For studies involving cancer, linkage of Medicare data with the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute is possible to obtain detailed information on cancer site, stage and histology. The Personal Summary File of Medicaid also contains state-specific case numbers identifying the Medicaid cases which each individual belongs to. Palmsten et al. (2013) used this number to identify mother-infant pairs. The authors showed that linkage is in general feasible, but the percentage of linked deliveries varies greatly by state.

1.4 Strengths and limitations

A major strength of Medicaid and Medicare is their enormous size, which enables studies of rare events even in small subpopulations. The induced homogeneity by the eligibility criteria increases control for confounding and restriction to subpopulations might be possible, where treatment effects might be detectable in contrast to the general population. However, the non-representativeness precludes studies aiming to describe the overall US population. Further, although a small proportion of patients is included in Medicaid without gaps of enrollment (Leonard et al. 2017), membership to Medicaid is generally not stable over time precluding studies of long-term effects of most treatments (Hennessy et al. 2012b). In contrast, members of Medicare generally stay in the program once they entered such that studies with long follow-ups are possible.

As with all administrative databases, Medicaid and Medicare lack information on important confounders such as lifestyle factors (e.g. smoking, diet) or occupation. Regarding prescriptions, the databases face the common problem that over-the-counter medication is not captured. Additionally, the prescription drug plans of Medicare Part D each differ with respect to drug coverage and cost-sharing options and drug availability thus differ across plans. Prescription data in Medicaid was shown to be accurate and complete (Leonard et al. 2017), but this might not hold for low cost generics (Choudhry and Shrank 2010).

1.5 Validation

Medicaid data of each state is validated internally by the CMS (Centers for Medicare & Medicaid Services 2016) and anomalies are reported in validation tables. For both Medicare and Medicaid, reimbursement of the health care providers is determined by the recorded procedures, which are thus checked for errors and can be considered as accurate. However, a validation study comparing medical records for surgical procedures of hip fractures with Medicaid claims found some procedures coded for another purpose (Wysowski and Baum 1993).

Outpatient prescription data of Medicaid was validated in the 1980s and found to be accurate (Lessler and Harris 1984). Leonard et al. (2017) recently investigated the quality of prescription claims in Medicaid and noted that 95-99% of the prescription claims were identifiable in a commercially-available database of NDCs. Further, the absolute number of prescriptions increased steadily and consistently over time, which suggest completeness of prescription claims. Validation of prescription claims was also performed with Medicare data. Colantonio et al. (2016) e.g. validated the prescription claims of lipid lowering drugs in Medicare with self-reported drug use and observed that many beneficiaries reported drug use although no claims in Medicare exist. However, both data sources might have caused this discrepancy e.g. due to recall error in self-reports or due to missing incentives to submit claims to Medicare for reimbursement.

Leonard et al. (2017) further found that hospital data in Medicaid was underrepresented in patients aged 45 years and above due to patients with poorer health status, who are dually eligible for Medicare and Medicaid. The authors therefore advised to additionally consider Medicare data for patients 45 years and older in case hospitalization data is needed to answer the study question of interest. Further, although no gross diagnostic miscoding in the in- and outpatient claims of Medicaid occurred, the authors acknowledged that the validity of health outcomes of interest generally remains open.

The gold-standard of outcome validation represents medical record validation, which was performed in a variety of studies. Hennessy et al. (2010) e.g. reviewed codes of inpatient and emergency department encounters in Medicaid to identify sudden cardiac death and ventricular arrhythmia originating in the outpatient setting and found very good agreement with medical records. Hernandez-Trujillo et al. (2015) performed medical record validation for primary immunodeficiency disease diagnoses. They observed low positive predicted values for individual ICD-9-CM codes and propose to use additional data sources to define disease status. Different methods to retrieve medical records were summarized in a Medicare-based example study validating adverse events of special interest (Wright et al. 2017). Other methods besides medical record validation were also performed. Brouwer et al. (2015) e.g. compared an algorithm to define myocardial infarction in Medicaid HIV patients with clinical cohort data and an algorithm to identify chronic kidney disease in older Medicare adults was validated with the Reasons for Geographic and Racial Differences in Stroke (REGARDS) study (Muntner et al. 2015).

1.6 Governance and ethical issues

The personnel information of Medicaid or Medicare is protected under the Privacy Rule of the Health Insurance Portability and Accountability Act of 1996, which prohibits disclosure of protected health information (PHI) without written consent. Exceptions are granted for

research purposes under certain conditions (Office for Civil Rights 2018). A study using PHI via research identifiable files of CMS requires a data use agreement and must be approved by the Privacy Board of CMS (Research Data Assistance Center 2020b, a). The study should assist CMS in monitoring, managing and improving the Medicaid and Medicare program and the services provided to the beneficiaries. Researchers must demonstrate experience in conducting research with files containing PHI and may only apply for the data files which are necessary to appropriately answer their study question. They further need to define a cohort in advance, since the provision costs depend on the number of patients in the cohort and the data files which are requested. However, there are some public use files containing non-identifiable data on a summary level, which can be requested without a data use agreement and without approval of the Privacy Board.

1.7 Documents and publications

The Research Data Assistance Center (ResDaC), a contractor funded by the CMS, provides introductory workshops and webinars on Medicaid, Medicare and the corresponding data files. The ResDac website also includes a detailed description of all information included in the RIFs. (<https://www.resdac.org/>).

Important methodological and applied example studies in pharmacoepidemiology were summarized by Hennessy et al. (2012b): Using routine data of elderly Medicaid beneficiaries of New Jersey Stürmer et al. (2005) compared conventional confounder adjustment with propensity score adjustment and adjustment using disease risk scores but found no major difference between the three methods. Schneeweiss et al. (2009) proposed an algorithm for automated confounder selection for propensity score models and evaluated its performance in three studies based on elder Medicare patients. The algorithm resulted in estimates that were all closer to the results of corresponding randomized clinical trials. Roumie et al. (2009) investigated the risk of cardiovascular events for certain non-steroidal anti-inflammatory drugs in Tennessee Medicaid enrollees and observed an increased risk for current users of rofecoxib, valdecoxib and indomethacin compared to non-users in patients without a history of cardiovascular diseases. The relation between adherence osteoporosis treatment and the risk of fractures in elderly Medicare beneficiaries was analyzed by Patrick et al. (2010), who found a consistent relation between adherence and risk reduction.

Recently conducted studies showed that Medicaid and Medicare data are still used frequently in diverse fields of epidemiological research: Leonard et al. (2018) compared new users of different antidiabetic monotherapies regarding the risk of severe hypoglycaemia in a cohort based on Medicaid beneficiaries from California, Florida, New York, Ohio and Pennsylvania and found the highest rate of serious hypoglycaemia for sulfonylureas. The prevalence of antidiabetic and antilipidemic medications in children and adolescents treated

with atypical antipsychotics was estimated by Varghese et al. (2016) in Virginia Medicaid beneficiaries. The authors observed that the medication was more often prescribed in atypical antipsychotic users than in non-users. Ray et al. (2015) compared time-dependent propensity scores with conventional adjustment of time-dependent confounders and inverse-probability-of-treatment (IPT) weighted estimation of parameters in marginal structural models in a cohort of opioid users of the Medicaid population in Tennessee in the absence of confounders on the causal pathway. IPT weighted estimates were shown to be less efficient than the other two in this example. Santos et al. (2016) analyzed the use of cytomegalovirus prophylaxis in kidney transplant recipients of Medicare and found that prophylaxis use was common. In an empirical example with Medicare patients, Gokhale et al. (2016) illustrated the considerable loss of power, when steps in the design study reduce sample size to minimize potential bias. Gilbertson et al. (2016) examined the influence of different time-windows for baseline confounder assessment on the mortality risk in a simulation study based on haemodialysis patients of Medicare and concluded that the timing of confounders should be taken into account for improvement of confounder control.

1.8 *Administrative information*

The administrative data of Medicaid and Medicare is maintained by CMS, but ResDaC helps with data requests.

Contact details

Organization/affiliation: Research Data Assistance Center
University of Minnesota School of Public Health
Division of Health Policy and Management
420 Delaware Street SE, Mayo D355
Minneapolis, MN 55455

Administrative Contact: resdac@umn.edu
1-888-973-7322

Website: <https://www.resdac.org/>

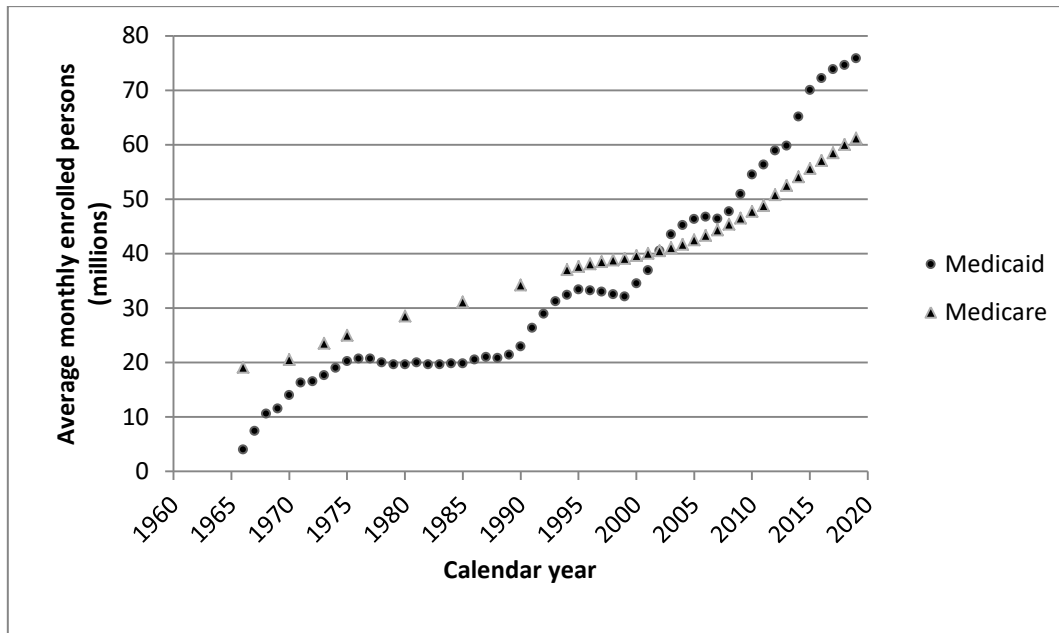


Figure 1: Enrolled persons in Medicaid (Medicaid and CHIP Payment and Access Commission 2019) and Medicare (Centers for Medicare & Medicaid Services 2019a, 2017, 2019b) from 1966 until 2016

Table 1: Demographic characteristics of the Medicaid and Medicare population

		Medicaid enrollees* (N=74,488,564)			Medicare enrollees (Part A and/or Part B)** (N=58,457,244)
Gender	Male	31,360,030 (42.10%)	Male		26,663,618 (45.61%)
	Female	43,035,479 (57.77%)	Female		31,793,626 (54.39%)
	Unknown	93,055 (0.12%)			
Age in years	<1	2,301,454 (3.09%)	Under 18		1,550 (0.00%)
	1-5	11,776,704 (15.81%)	18-24		96,951 (0.17%)
	6-12	13,353,457 (17.92%)	25-34		600,268 (1.03%)
	13-14	3,316,844 (4.45%)	35-44		1,123,996 (1.92%)
	15-18	6,428,030 (8.63%)	45-54		2,258,082 (3.86%)
	19-20	2,742,867 (3.68%)	55-64		4,698,364 (8.04%)
	21-44	18,930,113 (25.41%)	65-74		28,116,983 (48.10%)
	45-64	8,832,028 (11.86%)	75-84		14,944,692 (25.57%)
	65-74	3,092,551 (4.15%)	85-94		5,879,028 (10.06%)
	75-84	2,223,220 (2.98%)	95 and older		737,330 (1.26%)
	85+	1,406,428 (1.89%)			
	Missing	84,868 (0.11%)			
Race/Ethnicity	Non-Hispanic White	30,344,114 (40.74%)	Non-Hispanic White		43,452,430 (74.33%)
	Black/African American	15,771,904 (21.17%)	Black/African American		6,110,152 (10.45%)
	Hispanic/Latino	15,147,048 (20.33%)	Hispanic		5,292,863 (9.05%)
	Asian	2,496,906 (3.35%)	Asian/Pacific Islander		1,890,556 (3.23%)
	American Indian/Alaska Native	886,563 (1.19%)	American Indian/Alaska Native		269,595 (0.46%)
	Native Hawaiian/Pacific Islander	705,926 (0.95%)	Other		538,729 (0.92%)
	Hispanic/Latino and one or more race	3,360,206 (4.51%)	Unknown		902,918 (1.54%)
	More than one race	365,402 (0.49%)			
	Not identified	5,410,495 (7.26%)			

*Numbers based on Medicaid Statistical Information System (MSIS) tables of 2012 (Centers for Medicare & Medicaid Services 2012). Some states didn't provide numbers for the respective year and were thus extracted from the MSIS tables of 2011 (Arizona, Colorado, District of Columbia, Florida, Hawaii, Idaho, Kansas, Louisiana, Massachusetts, Texas, Utah) and 2010 (Maine)

**Numbers based on CMS Chronic Conditions Data Warehouse data on total enrollment in 2017 (Centers for Medicare & Medicaid Services 2019b).

References

- Brouwer ES, Napravnik S, Eron JJ, Jr. et al. (2015) Validation of Medicaid Claims-based Diagnosis of Myocardial Infarction Using an HIV Clinical Cohort. *Med Care* 53 (6):e41-e48. doi:10.1097/MLR.0b013e318287d6fd
- Centers for Medicare & Medicaid Services (2012) MSIS Tables. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MSIS-Tables.html>. Accessed 1st Feb 2018
- Centers for Medicare & Medicaid Services (2016) MAX Validation Reports. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX-Validation-Reports.html>. Accessed 04 Jan 2018
- Centers for Medicare & Medicaid Services (2017) Medicare and Medicaid Statistical Supplement 2013 Edition. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MedicareMedicaidStatSupp/2013.html>. Accessed 2 Feb 2020
- Centers for Medicare & Medicaid Services (2019a) CMS Fast Facts. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Fast-Facts/index.html>. Accessed 2 Feb 2020
- Centers for Medicare & Medicaid Services (2019b) Total Medicare Enrollment: Part A and/or Part B Enrollees, by Demographic Characteristics, Calendar Year 2015. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMSProgramStatistics/2015/2015_Enrollment.html. Accessed 2 Feb 2020
- Centers for Medicare & Medicaid Services (2020) Medicare-Medicaid Enrollment (MME): Total Medicare-Medicaid Enrollees by Type of Eligibility, Calendar Years 2010-2017. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMSProgramStatistics/2017/2017_Enrollment.html. Accessed 2 Feb 2020
- Choudhry NK, Shrank WH (2010) Four-Dollar Generics — Increased Accessibility, Impaired Quality Assurance. *N Engl J Med* 363 (20):1885-1887. doi:10.1056/NEJMp1006189
- Colantonio LD, Kent ST, Kilgore ML et al. (2016) Agreement between Medicare Pharmacy Claims, Self-report, and Medication Inventory for Assessing Lipid-Lowering Medication Use. *Pharmacoepidemiol Drug Saf* 25 (7):827-835. doi:10.1002/pds.3970
- Gilbertson DT, Bradbury BD, Wetmore JB et al. (2016) Controlling Confounding of Treatment Effects in Administrative Data in the Presence of Time-varying Baseline Confounders. *Pharmacoepidemiol Drug Saf* 25 (3):269-277. doi:10.1002/pds.3922
- Gokhale M, Buse JB, Pate V et al. (2016) More Realistic Power Estimation for New User, Active Comparator Studies: an Empirical Example. *Pharmacoepidemiol Drug Saf* 25 (4):462-466. doi:10.1002/pds.3872
- Hammill BG, Curtis LH, Qualls LG et al. (2015) Linkage of Laboratory Results to Medicare Fee-for-Service Claims. *Med Care* 53 (11):974-979. doi:10.1097/mlr.0000000000000420
- Hanlon JT, Donohue J (2010) Medicare Part D Data: A Valuable Tool for Pharmacoepidemiology and Pharmacoeconomic Research. *The American Journal of Geriatric Pharmacotherapy* 8 (6):483-484. doi:10.1016/S1543-5946(10)80001-7

- Hennessy S, Leonard CE, Freeman CP et al. (2010) Validation of Diagnostic Codes for Outpatient-Originating Sudden Cardiac Death and Ventricular Arrhythmia in Medicaid and Medicare Claims Data. *Pharmacoepidemiol Drug Saf* 19 (6):555-562. doi:10.1002/pds.1869
- Hennessy S, Palumbo Freeman C, Cunningham F (2012a) US Government Claims Databases. In: Strom BL, Kimmel SE, Hennessy S (eds) *Pharmacoepidemiology*. 5th edn. John Wiley & Sons, Chichester, pp 209-223
- Hennessy S, Palumbo Freeman C, Cunningham F (2012b) US Government Claims Databases. In: Strom BL, Kimmel SE, Hennessy S (eds) *Pharmacoepidemiology*. 5th edn. John Wiley & Sons, Chichester,
- Hernandez-Trujillo H, Orange JS, Roy JA et al. (2015) Validity of Primary Immunodeficiency Disease Diagnoses in United States Medicaid Data. *J Clin Immunol* 35 (6):566-572. doi:10.1007/s10875-015-0185-x
- Leonard CE, Brensinger CM, Nam Y et al. (2017) The Quality of Medicaid and Medicare Data Obtained from CMS and Its Contractors: Implications for Pharmacoepidemiology. *BMC Health Serv Res* 17 (304). doi:10.1186/s12913-017-2247-7
- Leonard CE, Han X, Brensinger CM et al. (2018) Comparative Risk of Serious Hypoglycemia with Oral Antidiabetic Monotherapy: a Retrospective Cohort Study. *Pharmacoepidemiol Drug Saf* 27 (1):9-18. doi:10.1002/pds.4337
- Lessler JT, Harris BSH (1984) Medicaid Data as a Source for Postmarketing Surveillance Information: Final Report. Research Triangle Institute, Research Triangle Park, NC
- Medicaid and CHIP Payment and Access Commission (2019) MACStats: Medicaid and CHIP Data Book. <https://www.macpac.gov/publication/macstats-medicaid-and-chip-data-book-2/>. Accessed 2 Feb 2020
- Muntner P, Gutierrez OM, Zhao H et al. (2015) Validation Study of Medicare Claims to Identify Older US Adults with CKD Using the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study. *Am J Kidney Dis* 65 (2):249-258. doi:10.1053/j.ajkd.2014.07.012
- Office for Civil Rights (2018) HIPAA Privacy. <https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html>. Accessed 2 Feb 2020
- Palmsten K, Huybrechts KF, Mogun H et al. (2013) Harnessing the Medicaid Analytic eXtract (MAX) to Evaluate Medications in Pregnancy: Design Considerations. *PLoS One* 8 (6):e67405. doi:10.1371/journal.pone.0067405
- Patrick AR, Brookhart MA, Losina E et al. (2010) The Complex Relation between Bisphosphonate Adherence and Fracture Reduction. *The Journal of Clinical Endocrinology and Metabolism* 95 (7):3251-3259. doi:10.1210/jc.2009-2778
- Ray WA, Liu Q, Shepherd BE (2015) Performance of Time-dependent Propensity Scores: a Pharmacoepidemiology Case Study. *Pharmacoepidemiol Drug Saf* 24 (1):98-106. doi:10.1002/pds.3727
- Research Data Assistance Center (2020a) Requirements for Institutional Review Board (IRB) Review and HIPAA Waiver Documentation for RIF DUA Request Submissions.

<https://www.resdac.org/articles/requirements-institutional-review-board-irb-review-and-hipaa-waiver-documentation-rif-dua>. Accessed 2 Feb 2020

Research Data Assistance Center (2020b) Research Identifiable Files (RIF) Requests. <https://www.resdac.org/cms-data/request/research-identifiable-files>. Accessed 2 Feb 2020

Roumie CL, Choma NN, Kaltenbach L et al. (2009) Non-aspirin NSAIDs, Cyclooxygenase-2 Inhibitors and Risk for Cardiovascular Events - Stroke, Acute Myocardial Infarction, and Death from Coronary Heart Disease. *Pharmacoepidemiol Drug Saf* 18 (11):1053-1063. doi:10.1002/pds.1820

Santos CA, Brennan DC, Saeed MJ et al. (2016) Pharmacoepidemiology of Cytomegalovirus Prophylaxis in a Large Retrospective Cohort of Kidney Transplant Recipients with Medicare Part D Coverage. *Clin Transplant* 30 (4):435-444. doi:10.1111/ctr.12706

Schneeweiss S, Rassen JA, Glynn RJ et al. (2009) High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data. *Epidemiology* 20 (4):512-522. doi:10.1097/EDE.0b013e3181a663cc

Stürmer T, Schneeweiss S, Brookhart MA et al. (2005) Analytic Strategies to Adjust Confounding Using Exposure Propensity Scores and Disease Risk Scores: Nonsteroidal Antiinflammatory Drugs and Short-term Mortality in the Elderly. *Am J Epidemiol* 161 (9):891-898. doi:10.1093/aje/kwi106

U.S. Census Bureau Population Division (2018) Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2018. <https://factfinder.census.gov>. Accessed 2 Feb 2020

Varghese D, Kirkwood CK, Carroll NV (2016) Prevalence of Antidiabetic and Antilipidemic Medications in Children and Adolescents Treated With Atypical Antipsychotics in a Virginia Medicaid Population. *Ann Pharmacother* 50 (6):463-470. doi:10.1177/1060028016638861

Wright NC, Delzell ES, Smith WK et al. (2017) Improving Medical Record Retrieval for Validation Studies in Medicare Data. *Pharmacoepidemiol Drug Saf* 26 (4):393-401. doi:10.1002/pds.4131

Wysowski DK, Baum C (1993) The Validity of Medicaid Diagnoses of Hip Fracture. *Am J Public Health* 83 (5):770