

Healthcare Payments Database (HPD) Data Completeness 2018-2023

Data completeness is an important aspect of data quality. A high percentage of data completeness does not measure accuracy of the submitted data, it represents the degree to which specific fields are populated with values under expected circumstances (numerator) across all records (denominator) for the expected circumstances.

Data is submitted to HPD monthly and at the service line level. Consequently, there is an eligibility record for every month a member was enrolled with a plan, and a service line record for every non-denied service line on a claim.

What Data is NOT in HPD

HPD system does not include the following information:

- Uninsured individuals
- Members covered through federal programs (such as Federal Employee Health Benefits Program, TRICARE, Veterans Affairs, and Indian Health Service)
- Data from exempt health plans, health insurers, and public self-insured plans below 40,000 members
- Private self-funded employers and Taft-Harley trusts regulated under the Employee Retirement Income Security Act (ERISA)

Private, self-insured companies interested in reducing costs and improving system performance are encouraged to participate in the HPD data collection on a voluntary basis.

What Data is in the HPD

The following types of payers submit data to HPD:

- Centers for Medicare and Medicaid Services (CMS) Medicare Fee-for-Service (FFS) members: Federal third-party reimbursement program administered by the Centers for Medicare and Medicaid Services that underwrites the medical costs of persons 65 and over, and some qualified persons under 65. Medicare FFS data is obtained from CMS through a state agency request process.

- Department of Health Care Services (DHCS) for Medi-Cal members (both FFS and managed care): A public health insurance program that provides free or low-cost medical services and healthcare benefits to low-income individuals, financed from state and federal funds; California’s version of Medicaid. Includes FFS Medi-Cal and managed care Medi-Cal.
- Commercial health plans and insurers for those with employer-based, individual, or Medicare Advantage coverage: Insurance products for which the coverage premium is paid by a private party, such as an employer, individual, or other entity.

CMS uses its own proprietary data format, but all other submitters use the All-Payers Database Common Data Layout ([Common Data Layout | APCD Council](#)). The APCD-CDL is comprised of the following file types:

- Eligibility
- Medical Claims
- Pharmacy Claims
- Provider

Discussion on FFS vs Managed Care Encounters

Encounter data captures information about healthcare services received when the provider does not require direct payment for those services. Such information includes data about the rendered service, diagnosis, patient, and provider. Managed care encounters capture services delivered to members enrolled in managed care plans that receive a capitated, per-member-per-month payment. FFS claims data captures billing and payment information for each individual service rendered by a provider. Both FFS and Managed Care Encounter data are included in the HPD dataset.

Discussion on Field Calculations

The following data provides an overview of completeness rates for HPD data collected between January 1, 2018, and December 31, 2023, broken down per payer type. Data on commercial plans is captured as a cumulative average of all records. Commercial range is provided to highlight variability in data submission rates across submitters. This document includes the most representative fields in each table and should not be regarded as an exhaustive list. Note that mandatory fields would have a higher completion rate due to a higher incentive of regulatory compliance. This analysis uses a record-level approach and evaluates completeness of the entire submitted entries in the HPD system.

Eligibility Data

The member eligibility file encompasses demographic information for each member eligible for medical and pharmacy benefits for one or more days of coverage at any time during the reporting period. Coverage type field identifies the member's coverage based on eligibility date and allows to distinguish between self-funded, short-term, and underwritten by the insurer plans. Product code field describes the member's type of insurance, as per the Accredited Standard Committee (ASC). Note that personally identifiable information (PII) fields such as member's first and last name, address, social security number, and birth date are excluded from standard limited datasets (SLD and SLD+). PII data is only authorized to eligible entities, such as university researchers, for the purposes of data linkage. Completion rates for race, ethnicity, and language are calculated as actionable fields, which exclude "unknown", "invalid", and "missing" values.

Field	Commercial Range	Commercial	Medicaid	Medicare FFS	Medicare Advantage
Total Records Available	N/A	2,079,481,159	2,119,192,710	349,652,231	538,126,258
Samples	0%-100%	76.0%	98.0%	99.0%	87.0%
Member ID (masked)	100%	100%	100%	100%	100%
Coverage Type Code	100%	100%	100%	100%	100%
Product Code	100%	100%	100%	100%	100%
First Name *	98.9%-100%	100%	100%	100%	100%
Last Name *	98.9%-100%	100%	100%	100%	100%
Physical Address *	97.1%-100%	98.6%	96.7%	0% **	98.7%
Zip Code (5)	99.3%-100%	99.9%	99.1%	100%	100%
Member Sex	100%	100%	100%	100%	100%
Member SSN *	0%-100%	94.3%	97.5%	100%	49.6
Date of Birth *	100%	100%	100%	100%	100%
Age in Years	99.7%-100%	99.9%	99.9%	100%	100%
Race (Actionable)	0%-99.4%	33.5%	89.7%	96.6%	49.8%

Field	Commercial Range	Commercial	Medicaid	Medicare FFS	Medicare Advantage
Ethnicity (Actionable)	0%-98.7%	48.5%	89.7%	0% **	36.2%
Language (Actionable)	0%-100%	66.2%	98.7%	0% **	82.4%

* Fields only available through an identifiable request

** Physical Address, Ethnicity, and Language are not provided by Medicare FFS

Medical Claims/Encounters Data

The medical claims file incorporates service-level claims and remittance information for medical services, including member demographics, provider details, payment amounts, clinical diagnosis codes, and procedure codes. The table below displays the data completion of key Medical claims and encounters metrics in the HPD system. The Commercial Range column below shows the minimum and maximum completion percentages of medical key metrics for Commercial submitters only. Fields that are unavailable for some payer types due to the nature of their service processing.

Field	Commercial Range	Commercial	Medi-Cal FFS	Medi-Cal Managed Care	Medicare FFS	Medicare Advantage
Total Records Available	N/A	1,950,813,845	866,567,577	1,908,712,324	1,831,466,122	842,626,164
Professional Records Available	N/A	1,517,661,940	409,363,992	1,345,168,557	1,013,889,434	575,771,429
Inpatient Records Available	N/A	48,667,985	29,490,782	94,368,423	149,137,939	45,065,512
FFS Records Available	N/A	859,136,542	866,567,577	N/A	1,831,466,122	190,725,086

Field	Commercial Range	Commercial	Medi-Cal FFS	Medi-Cal Managed Care	Medicare FFS	Medicare Advantage
Billing Provider NPI	75.9%-100%	97.2%	86%	100%	97.3%	97.2%
Procedure Code (Prof)	99.7%-100%	100%	100%	100%	100%	99.9%
Procedure Code Modifier (1)	16.0%-45.0%	24.8%	48.5%	34.1%	37.2%	25.2%
Date of Service	100%	100%	100%	100%	100%	100%
Rendering Provider NPI	67.6%-100%	97.1%	72.3%	100%	100%	97.3%
Payment Arrangement	100%	100%	100%	100%	100%	100%
Charge Amount (FFS)	86.9%-100%	99.9%	98.4%	N/A	100%	99.8%
Paid Amount (FFS)	93.7%-100%	99.9%	98.5%	N/A	100%	99.8%
Place of Service (Prof)	100%	100%	100%	100%	100%	100%
Type of Bill (Inpatient)	97.8%-100%	99.8%	100%	100%	100%	100%

Field	Commercial Range	Commercial	Medi-Cal FFS	Medi-Cal Managed Care	Medicare FFS	Medicare Advantage
Revenue Code (Inpatient)	99.8%-100%	100%	84.0%	100%	100%	100%
Principal Diagnosis	99.2%-100%	99.8%	87.0%	94.6%	100%	99.6%
Diagnosis Other (1)	0%-71.7%	54.9%	47.2%	58.2%	76.4%	62.9%

Pharmacy Claims/Encounters Data

The pharmacy claims file incorporates service-level claims and remittance information for prescription drug claims, including member demographics, provider details, payment amounts, and national drug codes (NDC). The table below displays the data completion of key pharmacy claims and encounters metrics in the HPD system. The Commercial Range column below shows the minimum and maximum completion percentages of pharmacy key metrics for commercial submitters only. Fields that are unavailable for some payer types due to the nature of their service processing.

Field	Commercial Range	Commercial Average	Medi-Cal FFS	Medi-Cal Managed Care	Medicare FFS	Medicare Advantage
Total Records Available	N/A	667,053,656	344,690,864	391,941,661	563,918,922	394,476,462
FFS Records Available	N/A	667,041,907	344,690,864	N/A	563,918,922	394,435,734
National Drug Code	100%	100%	99.8%	99.9%	100%	99.9%
Date Prescription Filled	100%	100%	100%	100%	100%	100%

Field	Commercial Range	Commercial Average	Medi-Cal FFS	Medi-Cal Managed Care	Medicare FFS	Medicare Advantage
Days Supply	99.5%-100%	99.9%	99.8%	100%	99.9%	99.9%
Payment Arrangement	100%	100%	100%	100%	100%	100%
Charge Amount (FFS)	92.5%-100%	99.6%	100%	N/A	99.7%	99.0%
Plan Paid Amount (FFS)	99.4%-100%	99.8%	100%	N/A	99.7%	99.1%
Pharmacy Provider NPI	93.3%-100%	99.9%	100%	100%	76.0%	100%
Prescribing Provider NPI	99.4%-100%	100%	99.2%	98.9%	100%	99.9%

Provider Data

Provider data is demographic and network information about providers associated with eligibility and claims data for a given reporting period. The completeness measurement covers the timeline from June 2017 to June 2024 and does not analyze the completion of fields by payer type in this fact sheet since the same provider can be present in data from multiple types of payers. Captured fields include provider identifiers, such as National Provider Identifier (NPI) and state license number, provider name and location, and provider type and primary specialty. HPD does not collect provider files from CMS; HPD cross-references provider data received from other sources to identify providers included on CMS Medicare FFS claims data.

Field	Completion Rates Range	Completion Rates
Total Records Available	N/A	121,829,559
Payer Assigned Provider ID	100%	100%
Entity Type Qualifier	0%-51.3%	13.1%
Provider NPI	90.1%-100%	97.2%
Provider State License Number	0%-96.8%	39.9%

Field	Completion Rates Range	Completion Rates
Provider First Name	9.2%-98.3%	59.7%
Provider Last Name or Organization Name	48.0%-99.6%	88.7%
Organization Name	0%-31.6%	2.3%
Provider Office Street Address	32.5%-98.4%	79.2%
Provider City	46.7%-98.9%	80.6%
Provider Office State	53.8%-98.9%	83.7%
Provider Office ZIP Code	53.9%-98.9%	83.6%
Provider Specialty	0%-87.4%	64.3%

References

- APCD Council. APCD-CDL Common Data Layout (version 3.0.1). Published April 1, 2023. Accessed February 17, 2025. [https://www.nahdo.org/sites/default/files/Resources/Data_Layouts/APCD-CDL v3.0.1_errata.pdf](https://www.nahdo.org/sites/default/files/Resources/Data_Layouts/APCD-CDL_v3.0.1_errata.pdf)
- California Department of Healthcare Access and Information. Health Care Payments Data Program, Report to the Legislature.; 2024. Accessed February 17, 2025. <https://hcai.ca.gov/wp-content/uploads/2024/02/HPD-Report-to-the-Legislature-March-2024-1.pdf>
- California Department of Healthcare Access and Information. Health Care Payments Data Program, Report to the Legislature.; 2020. Accessed February 17, 2025. <https://hcai.ca.gov/wp-content/uploads/2020/12/HPD-Legislative-Report-20200306-ADA.pdf>
- California Department of Healthcare Access and Information. Healthcare Payments Data (HPD) Services Report. Accessed February 17, 2025. <https://hcai.ca.gov/visualizations/healthcare-payments-data-hpd-services-report/>
- California Department of Healthcare Access and Information. Health Care Payments Data (HPD), Frequently Asked Questions. Accessed February 17, 2025. <https://hcai.ca.gov/data/cost-transparency/healthcare-payments/#faq>

Fact Sheet (Version 1.0) Updated March 19, 2025