Department of Health Care Access and Information

Program Report
Health Care Payments Data Program

March 2024

2020 WEST EL CAMINO AVENUE, SUITE 1222 • SACRAMENTO, CA 95833
PHONE: (916) 326-3600 • EMAIL: HCAIDO@HCAI.CA.GOV

"A healthier California where all receive equitable, affordable, and quality health care."

Table of Contents

Executive Summary

This Report provides a required program update to the California State Legislature on the Health Care Payments Data (HPD) Program, focusing specifically on the completeness and quality of data included in the initial five years of historical data that has been loaded in the HPD System.

Thanks to the coordinated efforts of many organizations and individuals, the HPD Program achieved several goals with the release of the first public data in June 2023, including:

- Established a new state program, modeled after recommendations made by the HPD Review Committee of stakeholders and experts (see the 2020 Report to the Legislature[1]), with infrastructure and processes to support continued updates and data access, pending ongoing state funding.[i]
- Met the legislative requirement to "substantially complete" the HPD System, the state's All-Payer Claims Database (APCD), by July 2023.
- Collected, aggregated, and loaded data from all planned sources in the state, including all Medi-Cal and Medicare Fee-for-Service covered lives, and all covered lives from California's health plans and insurers subject to the reporting mandate.
- Made ready for analysis five years of detailed healthcare service data for most of the Californian population, including member and utilization data for 82% of California's total population and 89% of California's insured population (for 2021).
- Published the first two sets of public data from the database:
    - HPD Snapshot[3]
    - HPD Measures[4]

As of the writing of this Report, the HPD System includes over 5 billion healthcare claim and encounter[ii] records from calendar years 2018 through 2022, with additional data received every month. Thirty-six submitting organizations representing California's health plans and insurers provide monthly submissions of eligibility, medical, prescription drug, and provider files. The HPD Program will continue to create annual extracts for reporting and analysis, growing the database over time. In addition, efforts to expand the types of data and submitters are already underway, including adding data from dental plans and insurers, capitation payments, and other non-claims payment data. Specifically, dental data collection will begin in 2024 and non-claims payment data is planned for 2025.

Preliminary analyses of data quality indicate that the quality is reflective of and consistent with administrative data used in healthcare operations—as well as other administrative data sources

[i] The Legislature provided $60 million in one-time funding via SB 840 (Mitchell, Chapter 29, Statutes of 2018) to establish the HPD Program. Spending authority for the initial $60 million expires at the end of June 2025. To support ongoing operations, HCAI recommends that state policy makers support an annual total funds budget of $22 million for the HPD Program, including $15.4 million in state funds, starting with Fiscal Year 2025-26 (see the 2023 Long-Term Funding Options Report[2]).

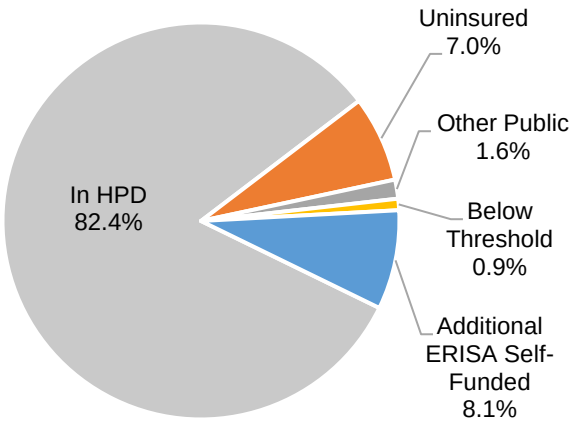
[ii] "Claims" refer to records of adjudicated fee-for-service claims between a provider and a plan; "encounters" are records similar to claims for services provided under a capitated payment arrangement. For more information on claims and encounters, see the subsection Background on Administrative Data in Health Care and Impacts to Data Quality.

and other state APCDs—and will support a wide range of research and analysis; and there are opportunities for improvement. In addition to continuing its work with submitters on data quality improvement efforts, HCAI will work with the Department of Managed Health Care and others to improve the collection, storage, and submission of standardized race, ethnicity, and language data.

Increasing submission of voluntary data from private self-funded employers and other purchasers provides the greatest opportunity to increase the comprehensiveness of the HPD. Due to a 2016 Supreme Court decision, private self-funded employers and Taft-Hartley trusts regulated under the Employee Retirement Income Security Act (ERISA) cannot be compelled to submit data to a state APCD, but they may do so voluntarily. Preliminary analysis of the self-funded lives in the HPD indicates that voluntary participation of ERISA plans is low and that as many as 3.2 million ERISA self-funded lives are not yet included in the HPD System. HCAI plans to meet with health plans, employers and other purchasers, and other stakeholders to discuss and implement strategies to increase the number of ERISA self-funded lives available in the HPD System.

Percent of Californians Represented in the HPD System, 2021

POPULATION GROUP	NUMBER	%
Included in the HPD	32,376,087	82.4%
Not Included in the HPD		
Uninsured	2,749,344	7.0%
Other Public (e.g., Military, Federal Employees, Indian Health Service)	609,000	1.6%
Below Threshold	365,428	0.9%
Additional ERISA Self-Funded	3,176,484	8.1%
Total Californians	**39,276,343**	**100%**



Notes:

- Number of uninsured from US Census Bureau.
- Number in Other Public from California Health Benefits Review Program, *Estimates of Sources of Health Insurance in California 2021*.
- Number below threshold based on HCAI analysis of covered lives reported in the California Health Care Foundation, California Health Insurers, Enrollment, 2023 Edition and HPD Program mandatory reporting thresholds. Includes regulated health plans and insurers only. A health plan, health insurer, or public self-insured plan that has fewer than 40,000 California members is not required to submit data to the HPD Program.
- Number in ERISA Self-Funded estimated from HCAI analysis and represents *additional* ERISA covered lives not already included in the HPD System. Derived by subtracting other categories from Total Californians. Note this may also include a small number of covered lives in public self-funded plans.
- Total Californians from US Census Bureau.

For calendar year 2021, the HPD System includes detailed healthcare services and enrollment data for 32.4 million average monthly members with coverage for medical care, or 82% of all Californians. Not included are the uninsured, those with coverage through federal programs

such as federal employees and the military, and some individuals covered by private self-funded employers and Taft-Hartley trusts.

California's APCD is as complete and representative of the state's population, or better, than other states' APCD programs, and is well positioned to meet the intent of the Legislature and the goals of the HPD Program to increase transparency in California's healthcare marketplace.

1. Introduction

In June 2018, the Governor signed Assembly Bill (AB) 1810 (Committee on Budget, Chapter 34, Statutes of 2018) which added Chapter 8.5, Health Care Cost Transparency Database, to the Health and Safety Code (HSC). Subsequently amended by AB 80 (Committee on Budget, Chapter 12, Statutes of 2020), HSC Sections 127671-127674 require the Department of Health Care Access and Information (HCAI) to plan for, develop, and administer a Healthcare Payments Data (HPD) System, often referred to as an All-Payer Claims Database (APCD) in the 19 states that have implemented such a program.[5]

This Report provides a program update to the California State Legislature on the HPD Program, focusing specifically on the amount and quality of the data included in the initial launch of the HPD System. HSC Section 127673, subdivision (k)(1) requires the report to include information on the following:

(A) Claims data reported by mandatory submitters.

(B) Claims data reported by voluntary submitters.

(C) Data on the covered lives reported, percentage of the population for which the department has data, the number of self-insured plans, providers and suppliers who have voluntarily submitted data, variation of completeness of data across geographic regions, such as the California Health Benefit Exchange's rating regions, the extent of data submitted on hospitals, physicians, and physician groups, the extent to which mandatory and voluntary submitters are submitting data specified in subdivision (b), frequency of submission of all core data, including claims, encounters, eligibility, and provider files, frequency of submission of nonclaims payment data files, and any other information that is available to determine if hospital and physician data are captured.

(D) A cost estimate if providers and suppliers become mandatory submitters.

(E) The number of data requests from qualified applicants and their data uses.

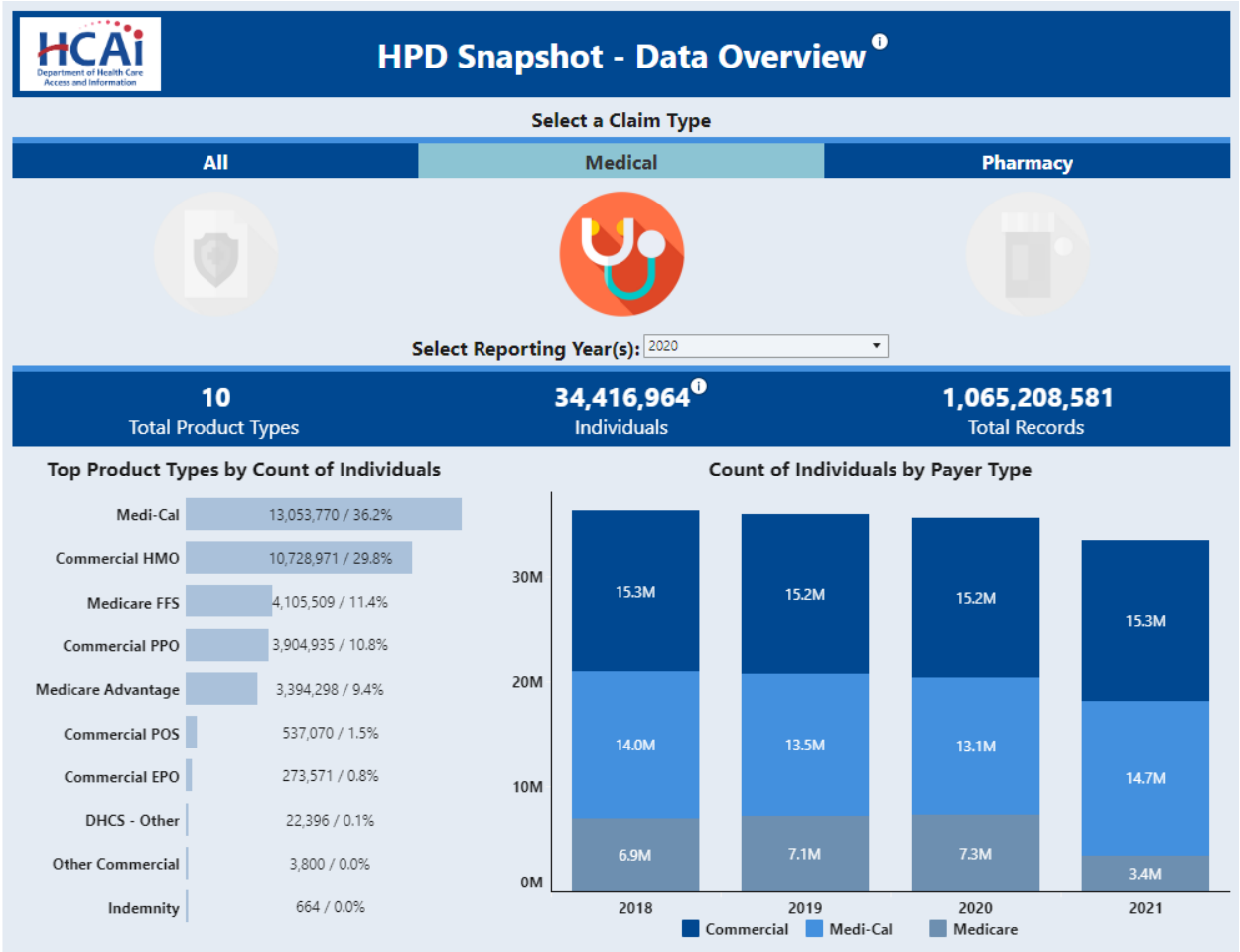
HCAI released the HPD Program's first public data in June 2023, including summary enrollment and healthcare utilization information for more than 30 million Californians, for calendar years 2018 through 2021. Publication of the HPD Snapshot[3] marked the successful culmination of a multi-year effort of legislation, planning, data collection, and implementation of California's APCD. The Snapshot provides an overview of data currently available as submitted in the HPD System with visualizations that allow users to explore how many Californians received coverage from each type of payer and the number of medical or pharmacy service records generated. Release of the Snapshot data also satisfied the legislative requirement that the development of the HPD System "be substantially completed" no later than July 1, 2023 (HSC §12671).

Exhibit 1 provides examples of the data and visualizations available in the Snapshot.

Exhibit 1. Screenshots from Publicly Available HPD Snapshot Visualizations

The Data Overview visualization shows the total number of records and how many Californians received coverage from each type of payer over time. Highlights from the example displayed below include:

- The HPD System includes over one billion medical records for 34.4 million individuals enrolled in medical coverage.
- For 2020, Medi-Cal members represented 36.2% of Californians in the HPD System.



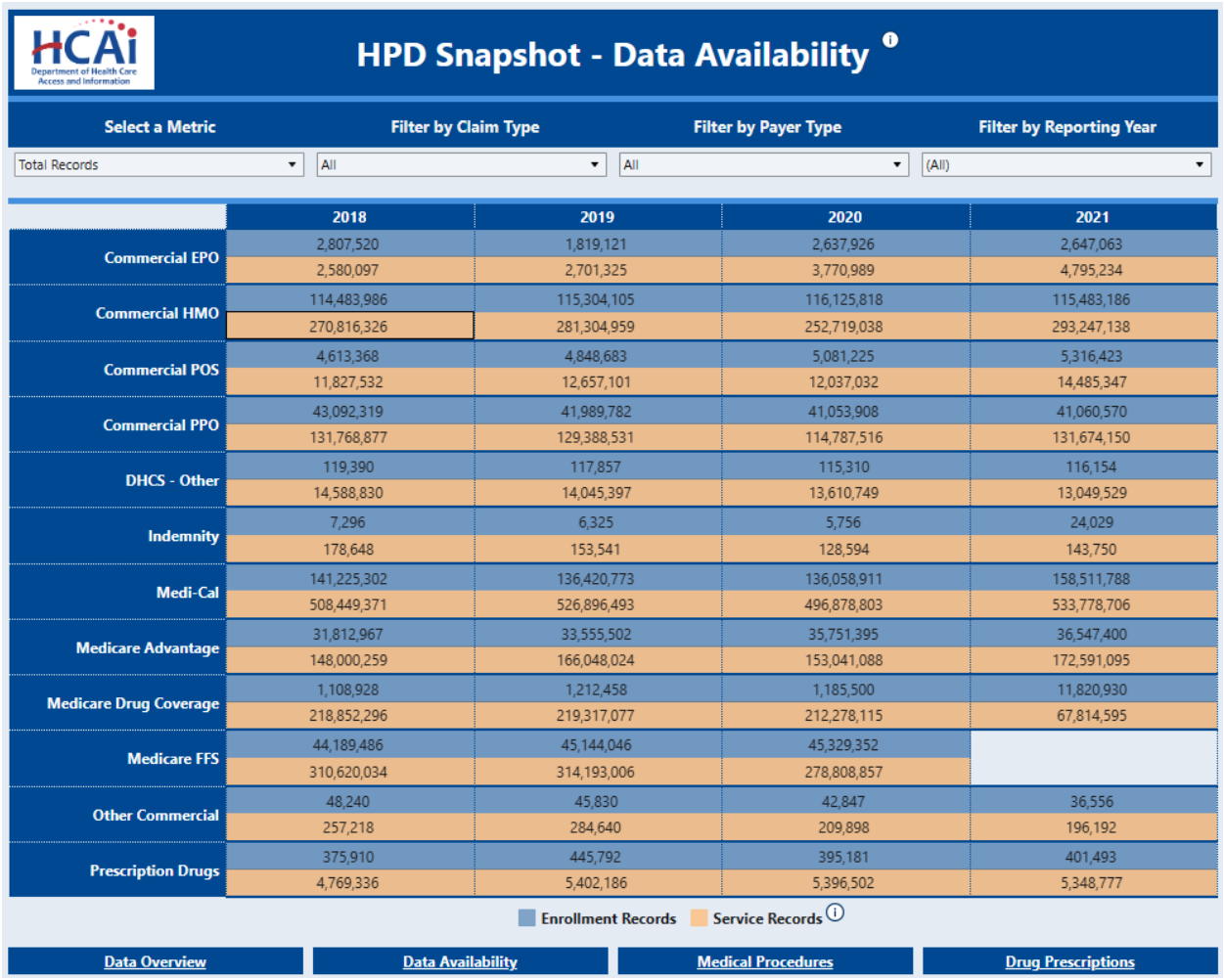
Notes:

- If an individual changes to a different product type or payer type within a calendar year, they will be included in the aggregate counts for both types that year but will only be counted once in the total number of covered individuals. Individuals enrolled in multiple product types at the same time will only be assigned to the product type that is identified as the primary payer. In cases where there is discrepancy in the reported data on which payer is primary, the order of assignment is: commercial payers, followed by Medicare, and then Medi-Cal.
- Medicare FFS data for 2021 were not available and are absent from the initial HPD Snapshot. Due to the handling of individuals enrolled in multiple product types at the same time described above, the count of individuals in the non-Medicare payer types is overstated for calendar year 2021.
- Additional information regarding how HCAI created this product is available at HPD Snapshot.

Exhibit 1, continued: Screenshots from Publicly Available HPD Snapshot Visualizations

The Data Availability visualization displays enrollment and service records available by type of product and coverage over time. Highlights from the example display include:

- Enrollment and service records are available for full calendar years 2018 through 2021 for all product types except for Medicare Fee-for-Service (FFS). Medicare FFS data for 2021 were not available and are absent from the initial HPD Snapshot.



HPD Snapshot - Data Availability

Select a Metric: Total Records | Filter by Claim Type: All | Filter by Payer Type: All | Filter by Reporting Year: (All)

	2018	2019	2020	2021
Commercial EPO	2,807,520	1,819,121	2,637,926	2,647,063
	2,580,097	2,701,325	3,770,989	4,795,234
Commercial HMO	114,483,986	115,304,105	116,125,818	115,483,186
	270,816,326	281,304,959	252,719,038	293,247,138
Commercial POS	4,613,368	4,848,683	5,081,225	5,316,423
	11,827,532	12,657,101	12,037,032	14,485,347
Commercial PPO	43,092,319	41,989,782	41,053,908	41,060,570
	131,768,877	129,388,531	114,787,516	131,674,150
DHCS - Other	119,390	117,857	115,310	116,154
	14,588,830	14,045,397	13,610,749	13,049,529
Indemnity	7,296	6,325	5,756	24,029
	178,648	153,541	128,594	143,750
Medi-Cal	141,225,302	136,420,773	136,058,911	158,511,788
	508,449,371	526,896,493	496,878,803	533,778,706
Medicare Advantage	31,812,967	33,555,502	35,751,395	36,547,400
	148,000,259	166,048,024	153,041,088	172,591,095
Medicare Drug Coverage	1,108,928	1,212,458	1,185,500	11,820,930
	218,852,296	219,317,077	212,278,115	67,814,595
Medicare FFS	44,189,486	45,144,046	45,329,352	
	310,620,034	314,193,006	278,808,857	
Other Commercial	48,240	45,830	42,847	36,556
	257,218	284,640	209,898	196,192
Prescription Drugs	375,910	445,792	395,181	401,493
	4,769,336	5,402,186	5,396,502	5,348,777

Enrollment Records | Service Records

Data Overview | Data Availability | Medical Procedures | Drug Prescriptions

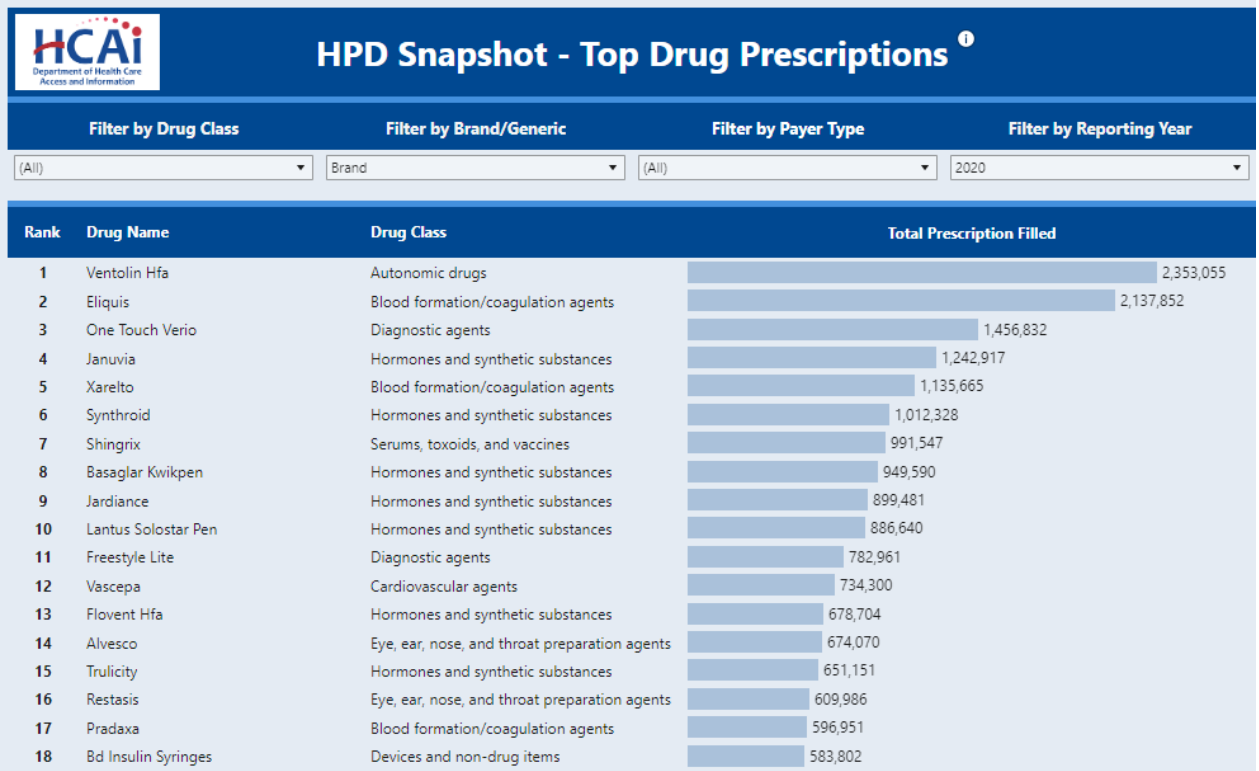
Notes:

- Information regarding how HCAI created this product is available at HPD Snapshot.

Exhibit 1, continued: Screenshots from Publicly Available HPD Snapshot Visualizations

The Top Drug Prescriptions visualization allows users to display the number of prescriptions filled, and filter by type of drug, payer type, and reporting year. Highlights from the example display include:

- The most commonly prescribed brand name drugs or supplies dispensed by a pharmacy in 2020 included:
  - Ventolin Hfa (albuterol sulfate), an oral inhaler for treatment or prevention of bronchospasm.[6]
  - Eliquis (apixaban), a factor Xa inhibitor anticoagulant indicated to reduce the risk of stroke.[7]
  - One Touch Verio, a blood glucose monitoring system.



Rank	Drug Name	Drug Class	Total Prescription Filled
1	Ventolin Hfa	Autonomic drugs	2,353,055
2	Eliquis	Blood formation/coagulation agents	2,137,852
3	One Touch Verio	Diagnostic agents	1,456,832
4	Januvia	Hormones and synthetic substances	1,242,917
5	Xarelto	Blood formation/coagulation agents	1,135,665
6	Synthroid	Hormones and synthetic substances	1,012,328
7	Shingrix	Serums, toxoids, and vaccines	991,547
8	Basaglar Kwikpen	Hormones and synthetic substances	949,590
9	Jardiance	Hormones and synthetic substances	899,481
10	Lantus Solostar Pen	Hormones and synthetic substances	886,640
11	Freestyle Lite	Diagnostic agents	782,961
12	Vascepa	Cardiovascular agents	734,300
13	Flovent Hfa	Hormones and synthetic substances	678,704
14	Alvesco	Eye, ear, nose, and throat preparation agents	674,070
15	Trulicity	Hormones and synthetic substances	651,151
16	Restasis	Eye, ear, nose, and throat preparation agents	609,986
17	Pradaxa	Blood formation/coagulation agents	596,951
18	Bd Insulin Syringes	Devices and non-drug items	583,802

Notes:

- Information regarding how HCAI created this product is available at HPD Snapshot.

In September 2023, HCAI released the HPD Measures[4] data visualizations. These visualizations allow users to explore the care and characteristics of Californians within the HPD System. The visualizations cover three measurement categories: Health conditions, Utilization, and Demographics. The health conditions measurements quantify the prevalence of long-term illnesses and major medical events, such as diabetes, asthma, and heart failure, in California's communities. Utilization measures present rates of healthcare system use through visits to the emergency department and different categories of inpatient stays, such as maternity or surgical stays. The demographic measures describe the health coverage and other characteristics (e.g., age group) of the Californians included in the data.

Each visualization presents the data in a different format to show geographic variation, changes over time, and comparisons to the statewide average. Filters and grouping options allow users to sort information by age group, sex, or location and to select specific populations. The combination of filtering options, visualization displays, and the collection of measures can answer a range of specific questions such as:

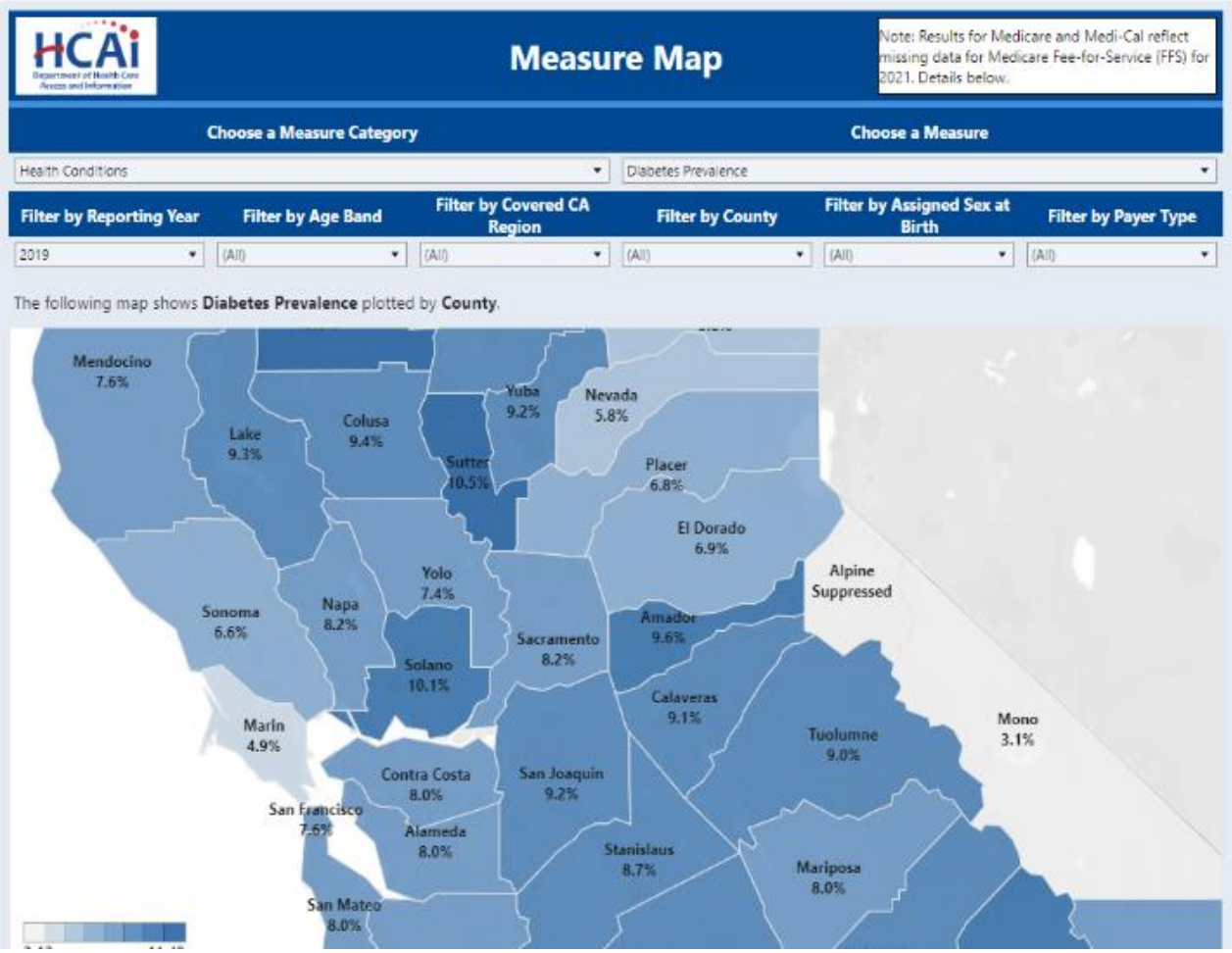
- What percentage of Californians in my age group have a diabetes diagnosis?
- Is the number of surgical inpatient stays increasing or decreasing over time?
- How does the share of the population enrolled in Medi-Cal in my county compare to the statewide average?

The visualizations include interactive data tables, trend charts, maps, and statewide comparison charts about health conditions, utilization, and demographics of populations represented in the HPD System. Exhibit 2 provides examples of the data and visualizations available in the Measures release.

Exhibit 2. Screenshots from Publicly Available HPD Measure Visualizations

The Measure Map allows users to display data for a variety of health conditions, utilization, and enrollment measures by county. Highlights from the example display include:

- Diabetes prevalence in Northern California counties ranged from a low of 3.1% in Mono County to 10.5% in Solano County (the state-wide average was 8.1%) in 2019.
- The rate for Alpine County is listed as "suppressed" because data from groups with small numbers are removed from the analyses, following the California Health and Human Services Agency's Data De-Identification Guidelines.[8]



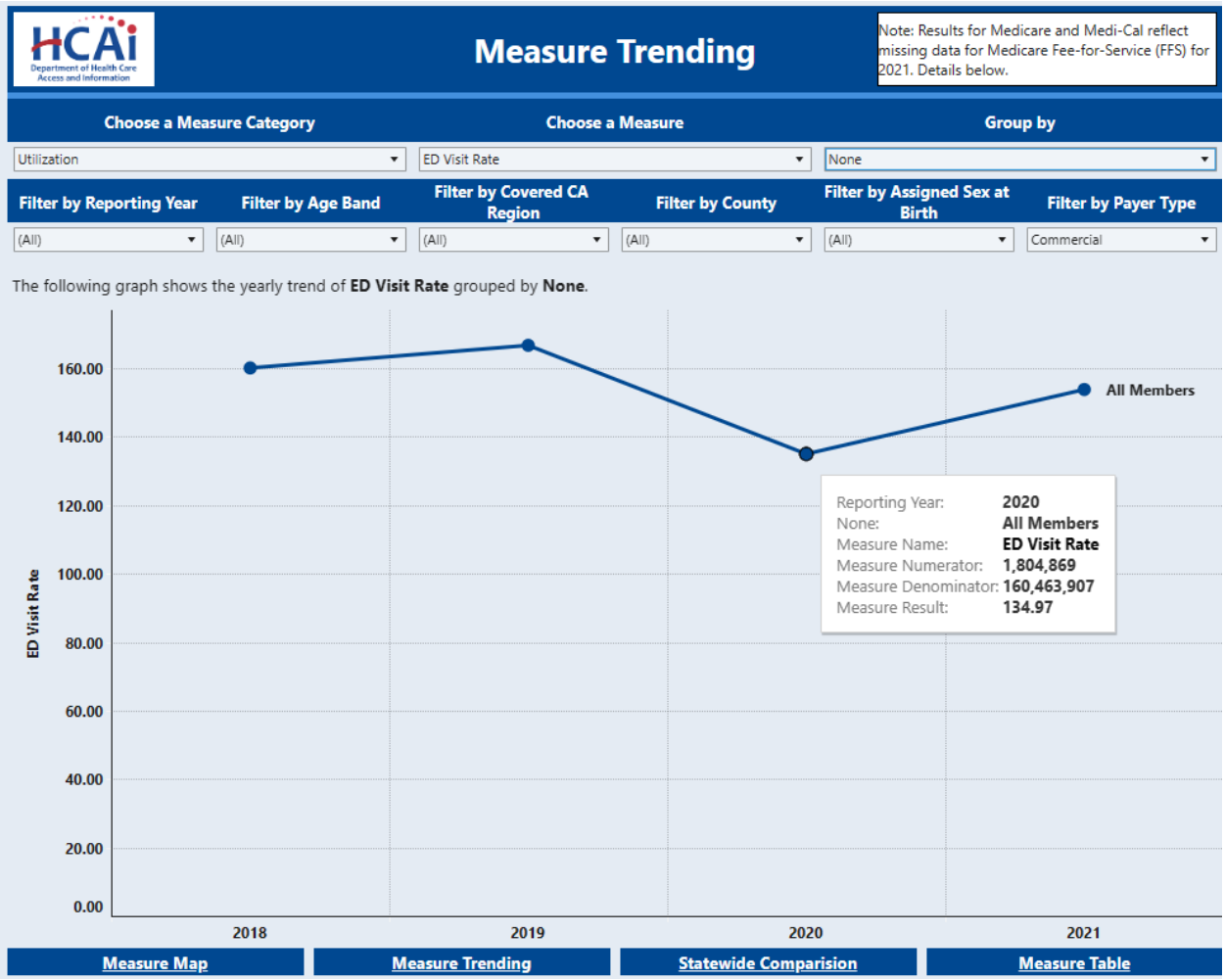
Notes:

- Information regarding how HCAI created this product is available at HPD Measures.

Exhibit 2, continued: Screenshots from Publicly Available HPD Measure Visualizations

The Measure Trending visualization allows users to display data for a variety of health conditions, utilization, and enrollment measures over time. Highlights from the example display include:

- The Emergency Department (ED) visit rate for members enrolled in commercial health plans declined considerably in 2020, to 135 visits per 1,000 members. Decreases in certain types of healthcare services occurred for a number of services during 2020, likely due to the COVID-19 pandemic.



Notes:

- Information regarding how HCAI created this product is available at HPD Measures.

2. Background on the HPD Program

With the passage AB 1810 in 2018, the California State Legislature established the Health Care Payments Data (HPD) Program, including the necessary planning, processes, resources, and system ("HPD System").[iii] In gathering, integrating, and organizing information about health plan and insurer payments for services, the HPD System offers an unprecedented opportunity to understand and address healthcare costs and drive improvement in California's healthcare system. The HPD Program will also play an important role in the Office of Health Care Affordability (OHCA). The California Health Care Quality and Affordability Act requires that OHCA use the HPD Program "to the greatest extent possible" to support the calculation of total healthcare expenditures (SB 184, Committee on Budget and Fiscal Review, Health, Chapter 47, Statutes of 2022).

HPD Program Goals[5]

1. Provide public benefit for Californians and the state while protecting individual privacy.
2. Increase transparency about healthcare costs, utilization, quality, and equity.
3. Inform policy decisions on topics including the provision of quality health care, improving public health, reducing disparities, advancing health coverage, reducing healthcare costs, and oversight of the healthcare system and healthcare companies.
4. Support the development of approaches, services and programs that deliver health care that is cost effective, responsive to the needs of Californians, and recognizes the diversity of California and the impacts of social determinants of health.
5. Support a sustainable healthcare system and more equitable access to affordable and quality health care for all.

Exhibit 3 displays the key program milestones, from initial legislation in 2018 to release of the first public reports in 2023.

Exhibit 3. Key HPD Program Milestones

DATE	MILESTONE	DESCRIPTION
Jun 2018	Initial Legislation – AB 1810 (Committee on Budget, Chapter 34, Statutes of 2018)	• Outlines HPD Program intent and requires planning effort • Requires the state to plan for, develop, and administer a "Health Care Cost Transparency Database," often referred to as an All-Payer Claims Database (APCD) in other states • Establishes the legislative intent of the HPD Program: – Establish a system to collect information regarding the cost of health care and a process for aggregating such information from many disparate systems, with the goal of providing greater transparency regarding healthcare costs

[iii] The following terms are used throughout this Report:

- "All-Payer Claims Database (APCD)" – the commonly used term across states for a large database of medical, dental, and pharmacy claims data, along with eligibility and provider files, from public and private payers within a state.
- "HPD Program" – the umbrella term for California's overall coordinated effort related to planning, implementing, and maintaining an APCD for the state of California, as outlined in HSC Chapter 8.5. Includes the organization, staffing, funding, processes, committees, workgroups, and other activities related to the effort.
- "HPD System" – term for California's APCD, including the underlying data and reports.

DATE	MILESTONE	DESCRIPTION
		– Improve data transparency to achieve a sustainable healthcare system with more equitable access to affordable and high-quality health care for all – Encourage use of such data to deliver health care that is cost effective and responsive to the needs of enrollees, including recognizing the diversity of California and the impact of social drivers of health
Mar 2019 - Feb 2020	HPD Review Committee	• Advises the state on the establishment, implementation, sustainability, and ongoing administration of the HPD Program
Mar 2020	Legislative Report[1]	• Includes background and learnings from other state APCDs, as well as 36 specific recommendations, discussed and voted on by Review Committee members, for the successful operation of the HPD Program in California, across nine areas: – APCDs and Use Cases – Data Categories and Formats – Linkages – Submitters – Funding and Sustainability – Privacy and Security – Technology Alternatives – Data Quality – Governance
Jul 2020	Updated Legislation – AB 80 (Committee on Budget, Chapter 12, Statutes of 2020)	• Authorizes data collection from health plans and insurers • Requires the establishment of an HPD Program Advisory Committee to assist and advise the state in formulating program policies regarding data collection, management, use, and access, and development of public information to meet the goals of the program • Identifies the types of data to be collected by the HPD Program, including detailed payment and healthcare utilization information • Requires the state to develop guidance (i.e., regulations) for required and voluntary data submissions from California's health plans, insurers, and other healthcare organizations • Specifies the make-up and requirements of a Data Release Committee to develop criteria, policies, and procedures for access to and release of data
Oct 2020	Advisory Committee	• Begins quarterly meetings • Serves as a forum for stakeholder and public engagement on policy decisions, while fostering accountability and transparency
Jul 2021	Submitter Group	• Begins quarterly meetings • Provides a forum for HPD data submitters to receive up to date information on submission requirements, troubleshoot data submissions, and address any other technical issues related to data submission
Dec 2021	Emergency Regulations	• Initiates the first stage of the HPD Data Program, including collecting core healthcare data, by identifying submitters, specifying data to be collected, creating a process for data submission, and establishing a timeline for data collection (Cal. Code Regs., tit 22, § 97300-97370)
May 2022	Plan and Submitter Registration	• Marks the completion of the registration process for California's required submitters – health plans and insurers identify key contacts, product offerings and attributes of plans, and relationships with other organizations responsible for submitting data

DATE	MILESTONE	DESCRIPTION
Jun 2022	Begin Initial Data Collection	• Submitters begin sending detailed data, including healthcare claims and encounters, eligibility, and provider information
Dec 2022	Data Release Committee	• Begins regular meetings to advise on criteria, policies, and procedures for access to and release of HPD data
Feb 2023	Complete Initial Data Collection	• Marks the completion of data delivery by submitters covering calendar years 2018-2021
Feb 2023	Legislative Report (Funding)[2]	• Summarizes long-term funding options for the program, for consideration by the legislature
Jun 2023	Public Reporting – HPD Snapshot	• Release of the first public data from the HPD, the Healthcare Payments Data (HPD) Snapshot, satisfying the legislative requirement that the development of the system "be substantially completed" no later than July 1, 2023 (HSC § 12671)
Sep 2023	Additional Public Reporting – HPD Measures	• Release of the second set of public data, Healthcare Payments Data (HPD) Measures. These visualizations allow users to explore the healthcare services, chronic conditions, and characteristics of Californians within the health system

3. What's Included in the HPD?

This section describes the data sources, submitters, volume of data, and comprehensiveness of the data included in the initial launch of the HPD System. The 2020 Health Care Payments Data Program Report to the Legislature included background and learnings from other state APCDs, as well as 36 specific recommendations, discussed and voted on by the HPD Review Committee.[1] These recommendations provided the planning foundation for the HPD Program, including the data types and sources for the initial implementation of the HPD System. The Report recommended three sources for data: 1) the California Department of Health Care Services (DHCS), for Medi-Cal Fee-For-Service (FFS) and Medi-Cal Managed Care; 2) the Centers for Medicare & Medicaid Services (CMS), for Medicare FFS; and 3) California's commercial plans and insurers, for their non-Medi-Cal members.

In June 2023, the HPD Program met its goal of successfully collecting and aggregating historical healthcare data for over 30 million Californians from each of the three major payer types within the state. Exhibit 4 compares the number of covered lives for each of those sources in the HPD System ("Actual") to the expected number of lives based on the planned data sources and mandatory data submission requirements of commercial plans and insurers ("Planned"). **The HPD System includes data from all the planned sources in the state, including all Medi-Cal and Medicare FFS covered lives and all covered lives from California's health plans and insurers subject to the reporting mandate**, including, for calendar year 2021:

- 14.1 million Medi-Cal members, including 2.4 million in FFS and 11.7 million in managed care plans
- 3.4 million members in Medicare FFS
- 16.8 million non-Medi-Cal members from California's health plans and insurers, including commercial and Medicare Advantage

Exhibit 4. Planned and Actual Covered Lives by Data Source, in Millions

SOURCE	2018	2019	2020	2021	2022
DHCS (Medi-Cal FFS)					
Planned	2.4	2.3	2.2	2.4	2.3
Actual	2.4	2.3	2.2	2.4	2.2
Percent	99.8%	99.9%	100.0%	99.4%	98.3%
DHCS (Medi-Cal Managed Care)					
Planned	10.8	10.5	10.7	11.7	12.8
Actual	10.8	10.5	10.7	11.7	12.8
Percent	99.9%	99.9%	99.9%	99.9%	99.9%
CMS (Medicare FFS)					
Planned	3.5	3.5	3.5	3.4	N/A
Actual	3.5	3.5	3.5	3.4	N/A
Percent	100.0%	>99.8%	100.0%	100.0%	N/A
California's Health Plans and Insurers (includes Medicare Advantage, excludes Medi-Cal)					
Planned	16.5	16.7	16.9	16.7	16.8
Actual	16.5	16.5	16.8	16.8	17.0
Percent	100.0%	98.9%	99.3%	100.8%	101.7%
Total (Unduplicated)					
Actual	31.5	31.3	31.6	32.4	N/A

Notes:

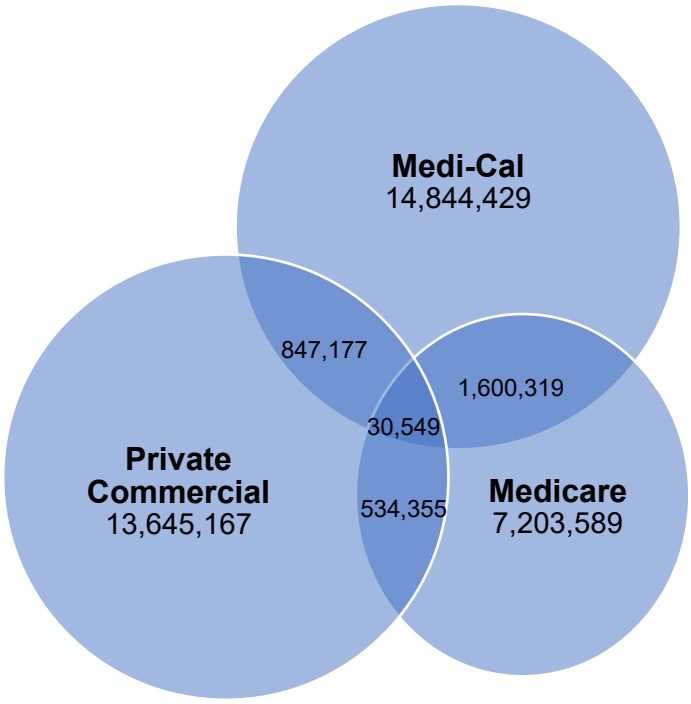
- Planned lives based on HCAI analysis of payers and types of data specified in HSC Sections 127671-127674 (i.e., DHCS, CMS for Medicare FFS, and Commercial Plans and Insurers with more than 40,000 non-Medi-Cal members for commercial and Medicare Advantage). Sources include Medi-Cal Certified Eligibles Tables (for DHCS), Kaiser Family Foundation State Health Facts (for CMS Medicare FFS), and California Health Care Foundation, 2023 Edition – California Health Insurers (for Commercial Plans and Insurers). Data for the following programs are included in the HPD System but excluded from the figures above if not enrolled in Medi-Cal: California Children's Services, Family Planning, Access, Care, and Treatment, and the Genetically Handicapped Persons Program.
- Actual lives based on HCAI analysis of average monthly covered lives in the HPD System.
- Actual lives for California's Health Plans and Insurers may be higher than Planned lives for 2021 and 2022 due to voluntary participation by self-funded employers. Please see Section 4, Opportunities to Expand the HPD, for more information on the number of self-funded lives in the HPD System.
- Medicare FFS data availability lags that of other sources; data for 2022 will be available in a subsequent release.
- The "Total (Unduplicated)" row represents the number of monthly individuals enrolled in one or more types of medical coverage by year. Unduplicated means that each individual is counted only once, even if they were enrolled in multiple types of coverage. The Unduplicated total is therefore not equal to the sum of the counts of individuals by source.

The HPD System, since it combines data from all the major payers in the state, provides the ability to quantify the number of individuals with more than one type of coverage between Medi-Cal, Medicare, and commercial plans. To illustrate, Exhibit 5 displays the number of individuals with multiple types of coverage for a single month (December 2021). Approximately 3.0 million out of the 32.7 million individuals in the HPD System had more than one type of medical coverage. In addition to 1.6 million Medi-Cal / Medicare dual members, there were 847,177 with Medi-Cal and private commercial coverage and 30,549 with coverage for all three types.

Exhibit 5. Distribution of Individuals by Type of Coverage, December 2021

COVERAGE TYPE	INDIVIDUALS	PERCENT
Medi-Cal Only	12,366,384	37.9%
Medicare Only	5,038,366	15.4%
Private Commercial Only (excludes Medicare Advantage and Medi-Cal)	12,233,086	37.5%
Medi-Cal & Private Commercial	847,177	2.6%
Medi-Cal & Medicare	1,600,319	4.9%
Medi-Cal & Private Commercial & Medicare	30,549	0.1%
Private Commercial & Medicare	534,355	1.6%
Total	**32,650,236**	**100.0%**

Multiple Coverage Types: 3,012,400 (9.2%)



HPD Program Data Sources

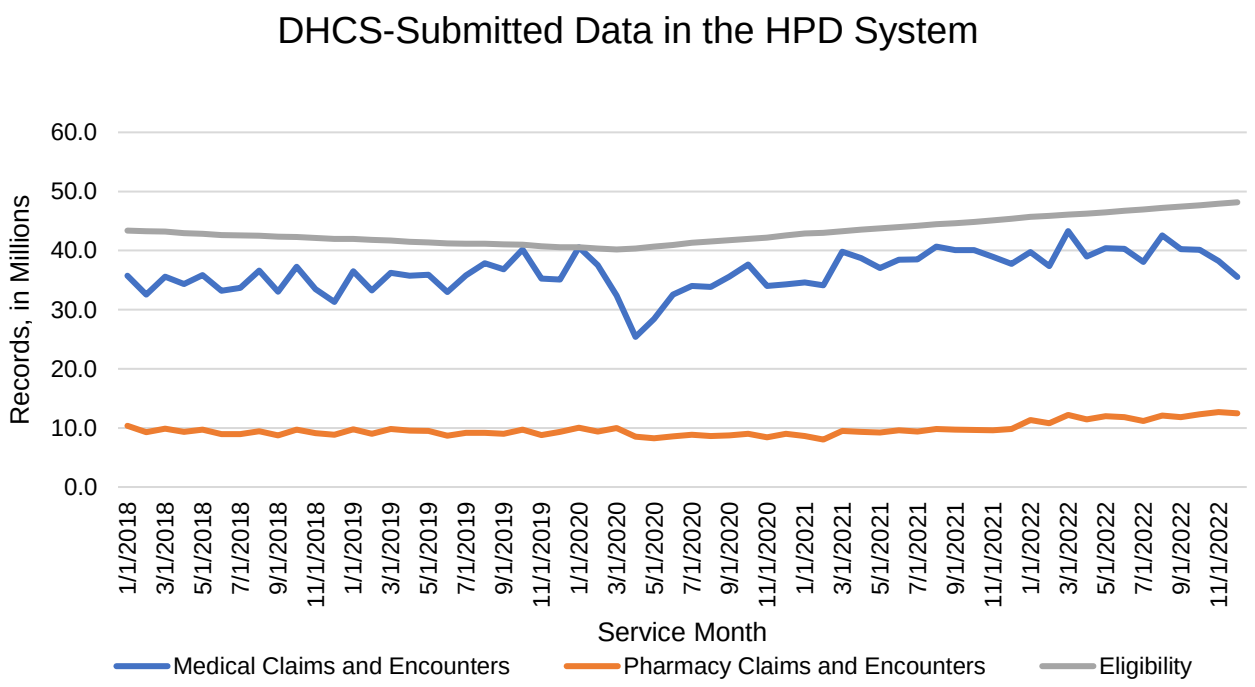
Department of Health Care Services (DHCS)

DHCS administers California's Medicaid program, Medi-Cal, which, in 2022, covered approximately 15 million members, 12.8 million through contracted managed care plans.[9] California residents eligible for Medi-Cal include low-income individuals and families, older adults, persons with disabilities, children in foster care, pregnant women, and low-income individuals with specific diseases. DHCS also administers and submits data to the

HPD Program for three other programs: California Family PACT (Planning, Access, Care and Treatment), California Children's Services, and the Genetically Handicapped Persons Program.

DHCS has a long history of collecting, storing, and analyzing detailed healthcare data for their programs. DHCS provides the HPD Program with monthly eligibility, medical claims and encounters, pharmacy claims and encounters, and provider data, following the Common Data Layout for State APCDs (APCD-CDL™), for both the Medi-Cal FFS and Medi-Cal Managed Care programs. The HPD System includes over 2.8 billion medical and pharmacy service records, and another 2.6 billion eligibility records for calendar years 2018-2022, from DHCS. Exhibit 6 displays the monthly record counts in the HPD System from DHCS. Note that the data show relatively steady numbers of eligibility and pharmacy records alongside a sharp decrease in medical services in spring 2020, consistent with the early months of the COVID-19 pandemic.

Exhibit 6. HPD System Records from DHCS



Centers for Medicare & Medicaid Services (CMS)

Medicare-eligible individuals are covered either through managed care (Medicare Advantage) or the traditional Medicare program, also known as Medicare FFS. In 2021, Medicare FFS covered approximately 3.4 million aged or disabled Californians for hospital and physician services; Medicare Advantage covered approximately 3.0 million.[10] The HPD Program obtains Medicare FFS and Prescription Drug Program data from CMS through a state agency request process, similar to most other state APCDs. The HPD Program collects California's Medicare Advantage data directly from California's health plans. Pharmacy benefits can be provided as part of a Medicare Advantage plan (Medicare Advantage - Prescription Drug Plan, or MA-PDP) or as a Stand-Alone Prescription Drug Plan to supplement the benefits in Medicare FFS or Medicare

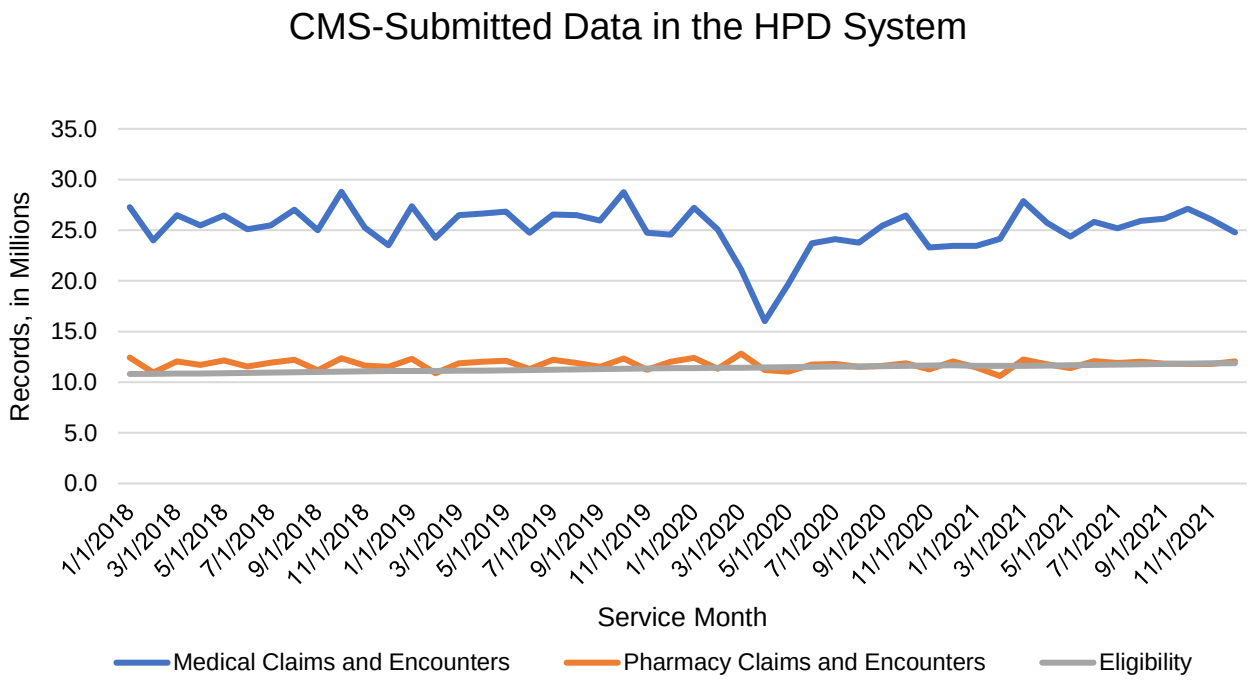
Advantage coverage. If pharmacy benefits are provided through a Medicare Advantage plan, data is submitted by the health plan; if pharmacy benefits are provided through a standalone PDP, data is acquired from CMS.

The HPD Program applies for and purchases the following Research Identifiable File types from the CMS Research Data Assistance Center (ResDAC) and integrates the Medicare FFS and Stand-Alone Prescription Drug Program data into the HPD:

- Master Beneficiary Summary File A/B/C/D
- Inpatient
- Outpatient
- Skilled Nursing Facility
- Hospice
- Home Health
- Carrier
- Durable Medical Equipment Regional Carrier (DMERC)
- Part D Event
- Part D Drug Characteristics
- Beneficiary Cross-walk files

The HPD System includes 1.8 billion medical and pharmacy service records and 545 million eligibility records from Medicare FFS and the Medicare Prescription Drug Program for calendar years 2018-2021 (due to the lag in CMS data availability, the 2022 data were not included in the HPD System at the time of this writing). Exhibit 7 displays the monthly record counts in the HPD System from CMS. Similar to the DHCS data, note that the data show steady enrollment alongside a sharp decrease in medical services in spring 2020, consistent with the early months of the COVID-19 pandemic.

Exhibit 7. HPD System Records from CMS



Mandatory Submitters

California Code of Regulations (22 CCR § 97310) related to the HPD Program specify the types of health plans and insurers that must submit data and the populations or types of coverage that must be included. Health plans and insurers regulated by the State of California that have at least 40,000 non-Medi-Cal covered lives (including Medicare Advantage) are considered mandatory submitters, as are Qualified Health Plans (Covered California plans), regardless of size. In addition to small plans, the regulations also exempt certain types of entities that are regulated by the federal Department of Labor, including employers and labor trusts (see also Section 4, Opportunities to Expand the HPD).

Participating commercial health plans and insurers provide the HPD Program with monthly eligibility, medical claims and encounters, pharmacy claims and encounters, and provider data, following the Common Data Layout for State APCDs (APCD-CDL™). These submitters are responsible for facilitating data submissions from appropriate data owners, including data feeds from different parts of the health plan or insurer, pharmacy benefit management companies, behavioral health organizations, subsidiaries, and other services carved out to a subcontracting organization. As a result, there are 48 submitting organizations representing plan entities that have registered with the HPD Program. The HPD System includes 3.5 billion medical and pharmacy service records and 2.2 billion eligibility records from these submitters for calendar years 2018-2022. Exhibit 8 displays the monthly record counts in the HPD System from these submitters. Similar to the DHCS and CMS data, note that the data show steady enrollment alongside a sharp decrease in medical services in spring 2020, consistent with the early months of the COVID-19 pandemic.

Exhibit 8. HPD System Records from Health Plans and Insurers

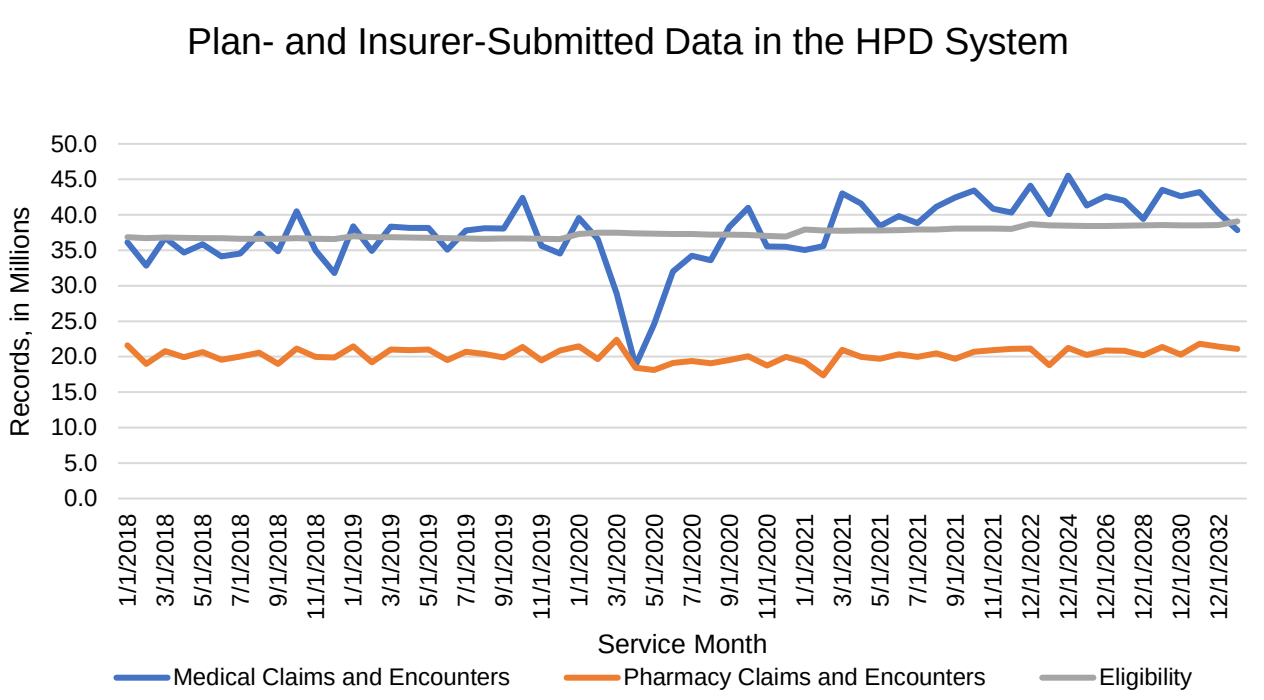


Exhibit 9 shows the number of planned vs. actual covered lives included in the HPD Program from commercial plans and insurers. The planned lives use the non-Medi-Cal enrollment data reported by health plans and insurers to the Department of Managed Health Care (DMHC) and the California Department of Insurance (CDI), as aggregated by the California Health Care Foundation,[11] applying the plan size threshold and other criteria included in the regulations for mandatory submitters. The data displayed in Exhibit 9 aggregates individually reporting and/or licensed plans and insurers into 19 “parent” plan rows for simplicity. Each row represents one or more licensed products or reporting entities, and each plan may also delegate data submission responsibility to additional organizations. Actual covered lives closely match planned covered lives for each year.

Of the mandatory plans, all but Universal Care (approximately 83,000 Medicare Advantage lives in 2022) successfully submitted all their data on time to be included in the current version of the HPD System. Universal Care has subsequently provided their data and it will be included in an upcoming release. **As of the writing of this Report, all mandatory plans and insurers have successfully complied with the data submission requirements of the HPD Program.**

Exhibit 9. Planned vs. Actual Covered Lives for California's Health Plans and Insurers (Note: includes Commercial and Medicare Advantage, excludes Medi-Cal)

	2018	2019	2020	2021	2022
Planned Lives by Parent Plan					
Kaiser*	7,991,311	8,153,577	8,215,029	8,215,896	8,171,657
Blue Shield*	2,710,362	2,652,350	2,554,819	2,352,122	2,329,359
Elevance (Anthem)*	2,148,692	2,249,953	2,238,963	2,097,673	2,209,675
UnitedHealth	1,225,351	1,176,900	1,245,967	1,355,824	1,365,606
Centene (Health Net)*	783,493	752,335	833,696	768,475	705,466
CVS (Aetna)	430,604	426,763	456,499	480,227	529,325
Cigna	329,565	345,717	343,656	330,258	308,210
SCAN	184,468	192,143	205,819	206,870	257,630
Sharp*	142,649	134,753	144,603	141,651	148,153
Humana	91,316	105,840	120,137	130,045	128,209
L.A. Care*	68,181	73,646	75,849	101,822	112,720
Western Health Advantage*	125,882	127,851	101,791	102,677	103,460
Sutter	83,874	93,987	96,692	100,466	102,768
Alignment	40,309	49,313	68,323	83,509	92,994
Universal Care	<threshold	<threshold	55,064	69,706	83,036
Molina	50,208	50,467	46,431	63,598	66,101
Oscar*	39,609	55,894	103,833	96,831	62,457
Valley Health Plan*	34,042	39,268	32,907	33,119	44,709
Chinese Community Health Plan*	20,444	18,738	13,815	12,688	11,738
Total Planned Lives	**16,500,360**	**16,699,495**	**16,898,829**	**16,673,751**	**16,750,237**
Total Actual Lives in HPD System	**16,496,964**	**16,511,883**	**16,788,320**	**16,814,321**	**17,028,733**
Percent	**>99.9%**	**98.9%**	**99.3%**	**100.8%**	**101.7%**

Notes:

- Planned lives based on HCAI analysis of mandatory health plan and insurer submitters specified in HSC Sections 127671-127674 and enrollment data from the California Health Care Foundation, 2023 Edition – California Health Insurers. A health plan, health insurer, or public self-insured plan that has fewer than 40,000 California members is not required to submit data to the HPD Program. Covered California plans must submit data regardless of plan size.
- Actual lives based on HCAI analysis of average monthly covered lives in the HPD System and may be higher than Planned lives due to voluntary participation by self-funded employers.
- * = Covered California plan (mandatory submitter regardless of covered lives).

Voluntary Submitters

Health plans that don't meet the definition of a mandatory submitter may choose to submit their data voluntarily. These include:

- Plans sponsored by self-funded employers and other purchasers that are not regulated by the State of California.
- Health plans and insurers below the 40,000 covered life threshold (the threshold is applied to each licensed entity) that are not Covered California plans.

Self-funded employers that are not mandated to submit data to the HPD can work with their plan administrators to voluntarily provide data on their behalf; some of that data is already in the HPD System. California Schools VEBA, a private joint labor management trust that provides health care for more than 65 participating education, municipal and public agency employers,[12] is actively working with HCAI to voluntarily submit data for its self-funded plan options. The HPD System also includes data from University of California self-funded plan options that are below the 40,000 threshold. Please see Section 4, Opportunities to Expand the HPD, for more information on the number of self-funded lives in the HPD System.

What Proportion of California's Population and Healthcare Services are Represented in the HPD System?

The following sections compare various data represented in the HPD System to external sources of similar data, including covered lives, ED visits, inpatient discharges, office visits, and prescription drugs. These sections also provide information on the number of providers represented in the HPD. Note that the external sources are intended to provide a general comparison and an initial evaluation of the amount and extent of data in the HPD System—differences in the included populations and methodologies between the HPD System and the comparison sources prevents a perfect comparison.

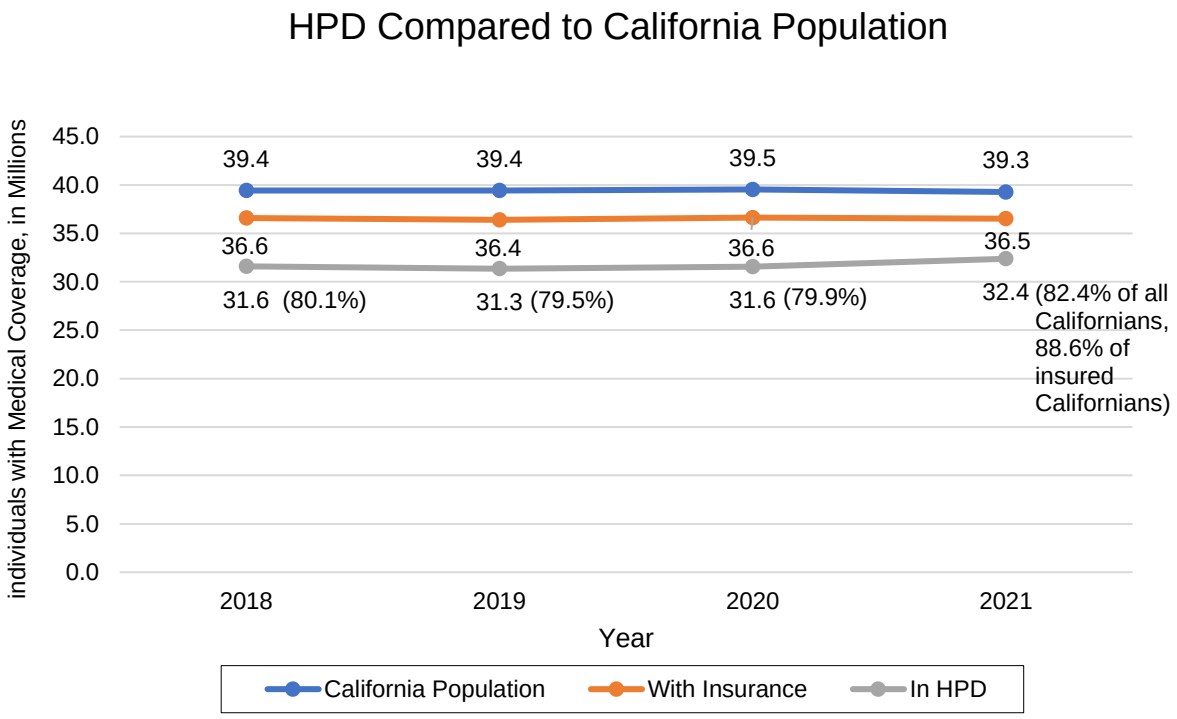
The comparisons presented below indicate that the HPD System represents approximately 80% of California's healthcare experience. Specifically, the HPD System includes approximately:

- Member information for 82% of California's total population and 89% of California's insured population
- 90% of state-wide ED visits
- 85% of inpatient admissions
- 76-89% of office visits

Covered Lives

As shown in Exhibit 10, the HPD System includes services and eligibility records for more than 31 million individuals with medical coverage in one or more plans per year, **representing approximately 82% of California's population in 2021**.[13] After adjusting for those without insurance (approximately 7%)[14], the HPD includes detailed healthcare service data for approximately 89% of insured Californians.

Exhibit 10. Covered Lives in the HPD Compared to California's Total and Insured Population



Notes:

- HPD covered lives reflect those with medical coverage.
- California population estimates and rate of uninsured from US Census Bureau.
- See Section 4, Opportunities to Expand the HPD, for information about efforts to increase the number of covered lives in the HPD System.

The percentage of each county's population reflected in the HPD is displayed in Exhibit 11. As expected from the 82% state-wide total, the HPD data for most counties are between 75% and 85%. Possible causes of variance by county include county-specific rates of uninsured and the number of individuals enrolled in Employee Retirement Income Security Act (ERISA) self-funded plans not included in the HPD System. Please see Section 4, Opportunities to Expand the HPD, for more information on ERISA self-funded plan participation.

Exhibit 11. Percent of Californians in the HPD System, by County, 2021

PERCENT OF 2021 POPULATION IN HPD, COUNTY	PERCENT OF 2021 POPULATION IN HPD, COUNTY
Counties with 59-69% of Population in HPD	**Counties with 80-84% of Population in HPD**
Alpine	Imperial
Lassen	Stanislaus
Mono	Kern
San Benito	Tulare
Sierra	Fresno
Trinity	Humboldt
Counties with 70-74% of Population in HPD	Lake

PERCENT OF 2021 POPULATION IN HPD, COUNTY	PERCENT OF 2021 POPULATION IN HPD, COUNTY
Amador	Santa Cruz
Kings	Madera
Mariposa	San Francisco
Modoc	Shasta
Plumas	San Bernardino
San Diego	Marin
Tehama	Alameda
Tuolumne	Mendocino
Yuba	Solano
Counties with 75-79% of Population in HPD	Merced
Calaveras	San Joaquin
El Dorado	Butte
Monterey	Del Norte
Nevada	Contra Costa
Orange	**Counties with 85-92% of Population in HPD**
San Luis Obispo	Colusa
Sutter	Glenn
Placer	Inyo
Ventura	Los Angeles
San Mateo	Napa
Santa Barbara	Sacramento
Riverside	Siskiyou
Santa Clara	Sonoma
Yolo	

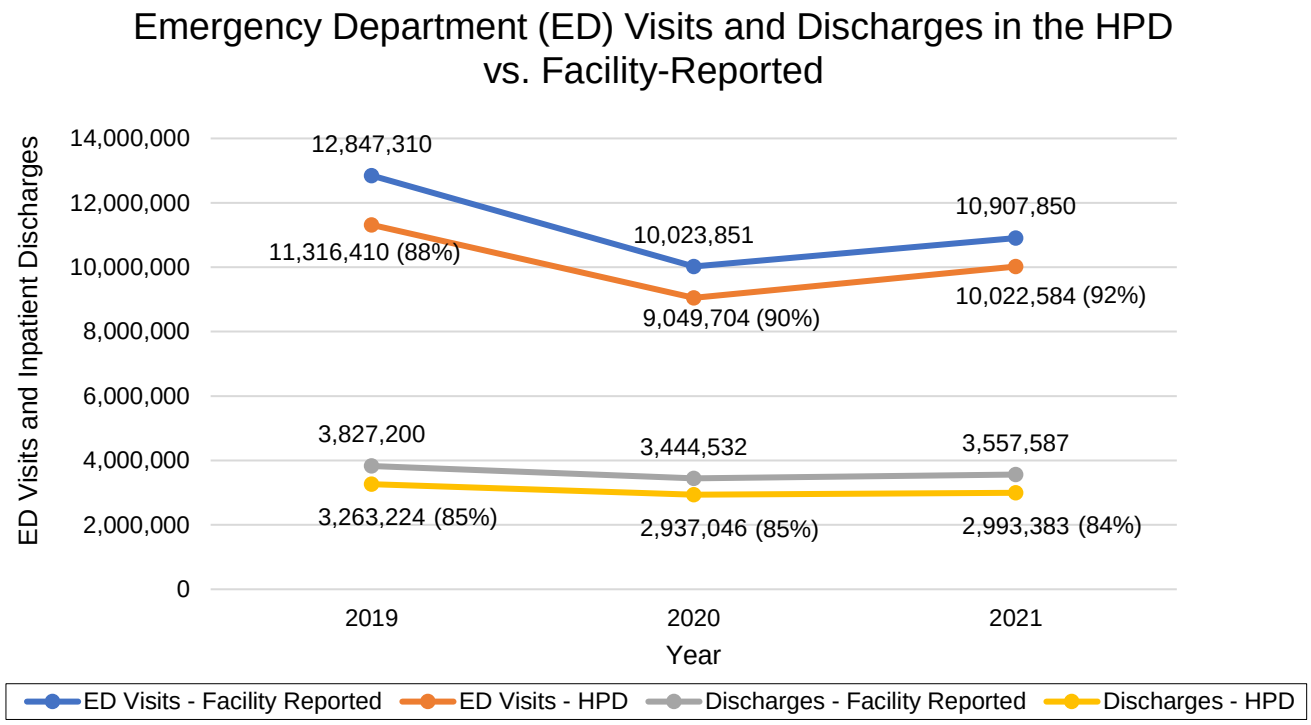
Notes:
- Population by county from US Census Bureau, Annual Estimates of the Resident Population for Counties in California: April 1, 2020 to July 1, 2022 (CO-EST2022-POP-06).

ED Visits and Inpatient Discharges

Separate from the HPD Program's data collection efforts, HCAI also collects and manages data from over 9,500 California licensed healthcare facilities and other healthcare entities and maintains several healthcare data reporting programs. Since all licensed facilities report their data, regardless of insurance status of the patient, these facility-provided data provide a benchmark for the potential universe of healthcare services provided by facilities in California. Exhibit 12 compares the numbers of ED visits and inpatient discharges in the HPD System to data separately reported by facilities to HCAI. For calendar year 2021, the HPD includes 10.0 million ED visits and 3.0 million inpatient discharges, up slightly from 2020 but down from 2019. Consistent with the comparisons to the total population of California, the ED visits and inpatient discharges in the HPD are slightly lower than the state-wide totals—the HPD data for 2021

includes approximately 92% of the ED visits and 84% of the inpatient discharges reported separately to HCAI.

Exhibit 12. ED Visits and Inpatient Discharges in the HPD Compared to HCAI Patient Discharge Data and ED Data



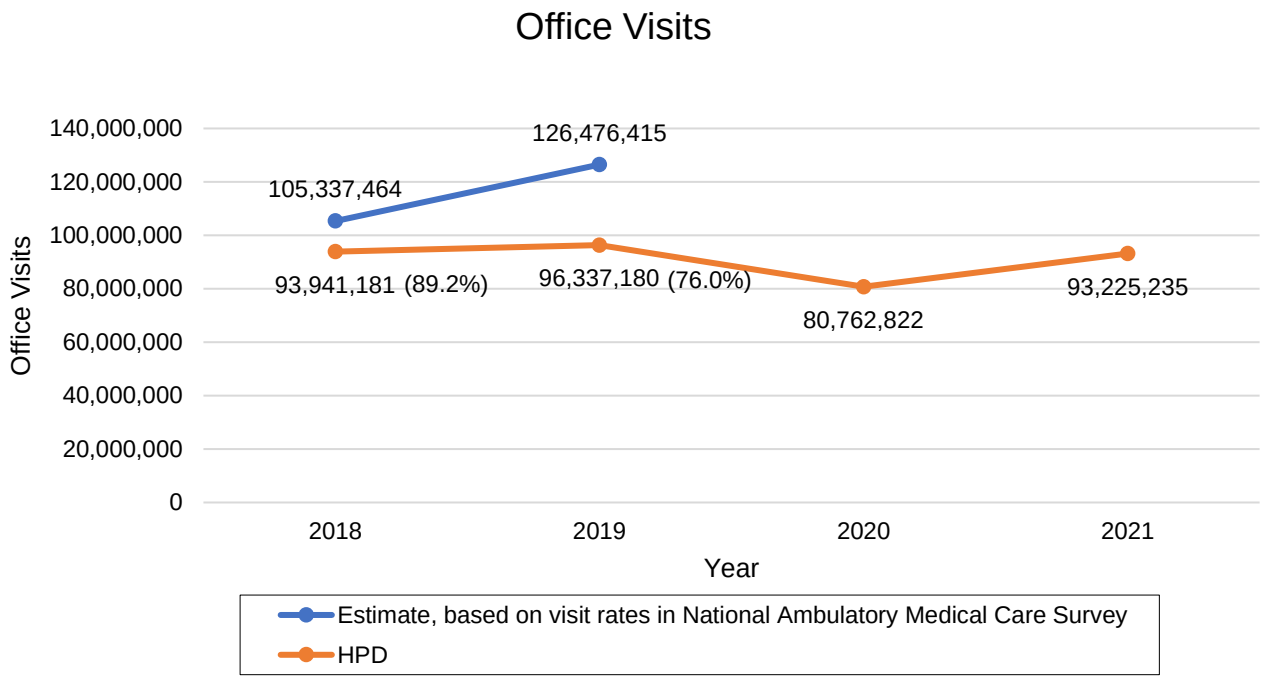
Notes:

- ED Visits - Facility Reported from HCAI, Hospital Emergency Department - Encounters by Facility.
- Discharges - Facility Reported from HCAI, Hospital Inpatient - Characteristics by Facility (Pivot Profile).
- Differences in the methodologies between the HPD System and the comparison sources may contribute to the variances displayed in the exhibit.

Office Visits

Exhibit 13 shows the number of office visits to a provider that are currently available in the HPD, and, for 2018 and 2019, estimates of visits statewide based on national survey data from the National Ambulatory Medical Care Survey. Although not state-specific and based on survey data (vs. claims data), the estimate provides a comparison point. If California office visit rates are the same as the national average, the HPD represents approximately 76%-89% of all office visits to a provider in the state.

Exhibit 13. Office Visits in the HPD Compared to Estimates from Survey Data



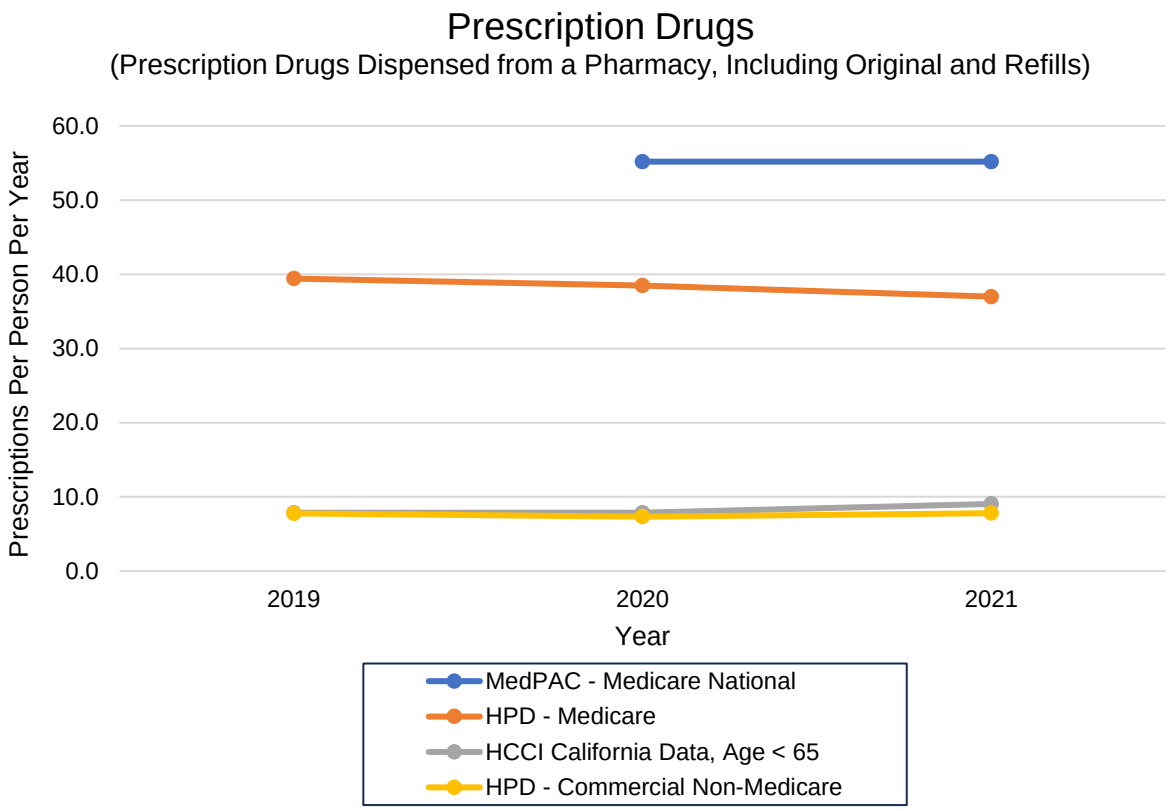
Notes:

- HPD Data from June 2023 HPD Snapshot, using Procedure Category of "Office/outpatient services - Office visits" and Type of Setting "Provider."
- Comparison data from the National Center for Health Statistics (NHCS) National Ambulatory Medical Care Survey (NAMCS): 2018 National Summary Tables and 2019 National Summary Tables, Office Visits to a Physician, scaled for California's population size. NAMCS data not available for 2020 and 2021. NHCS notes the 20% increase between 2018 and 2019 and that caution should be used when comparing 2018 and 2019 NAMCS estimates.
- Differences in the included populations and methodologies between the HPD System and the comparison sources may contribute to the variances displayed in the exhibit.

Prescription Drugs

Exhibit 14 shows the number of prescriptions per person per year in the HPD System for commercial and Medicare members compared to external sources. For non-Medicare members in commercial plans, prescriptions per member in the HPD range from 7-8 per year, which is comparable to the 7-9 range in data published by the Health Care Cost Institute. For Medicare members, the HPD System shows approximately 37-40 prescriptions per year, lower than the national average for Medicare of 55 prescriptions per individual per year.

Exhibit 14. Prescription Drugs in the HPD Compared to External Data



Notes:

- HPD data from pharmacy records and includes initial prescriptions and refills for drugs and supplies. Each prescription is counted once regardless of the number of days supply.
- HCCI data from Health Care Cost Institute (HCCI), Health Care Cost and Utilization Report, representing California data for individuals under age 65 with employer-sponsored insurance.
- MedPAC data from The Medicare Payment Advisory Commission (MedPAC) 2022 and 2023 Data Books, Health Care Spending and the Medicare Program, Chart 10-17. Prescriptions are standardized to a 30-day supply.
- Differences in the included populations and methodologies between the HPD System and the comparison sources may contribute to the variances displayed in the exhibit.

Hospitals and Physicians

The files collected by the HPD Program include several data elements that allow the identification of the providers that delivered and billed for services, including hospitals and other facility types, physician group practices, and individual physicians. Key elements include the provider name, address, type, and provider identifiers. The HPD System also incorporates information from the National Plan and Provider Enumeration System, which supports the inclusion of additional data elements including taxonomy code that describes the provider or organization's type, classification, and the area of specialization.[15]

Based on the comparisons to external data, most of California's providers are represented in the HPD System. Such provider data can be used to support analyses at the appropriate level of granularity and that is credible and supportable with the available data. As with all APCDs, there are challenges related to provider analyses. For example, despite the implementation of the National Provider Identifier in 2008,[16] a single hospital may have several "sub-parts" for billing

purposes that must be aggregated in analyses of the data. There are also challenges related to consistent use of billing and rendering provider identifiers in administrative data.

Exhibit 15 compares the number of discharges reported to HCAI pursuant to HSC Section 128735, subdivision (g), and compares these facility-reported discharges to the number of discharges calculated from the administrative data in the HPD System. For calendar year 2021, total discharges in the HPD System are 84% of those reported separately by facilities to HCAI, reflecting that the HPD System is capturing a large majority of inpatient discharges statewide. However, the share of admissions to the ten facilities with the highest number of discharges (representing approximately 10% of all discharges) is somewhat lower, at 65%. HCAI continues to evaluate the data, reasons for the variance, and opportunities to increase the share of inpatient facilities represented in the HPD System. The comparison highlights challenges related to reliably identifying inpatient stays and the admitting facility in administrative data. As the HPD System matures, these comparisons are anticipated to improve over time. See also Known Data Qualities Issues in Section 5.

Exhibit 15. Discharges in the HCAI Patient Discharge Data vs. HPD, 2021

RANK	HCAI PATIENT DISCHARGE DATA	HPD	PERCENT
Top Ten Facilities in HCAI Patient Discharge Data, Ranked by Number of Discharges	371,605	242,676	65%
All Other	3,185,982	2,750,707	86%
Total	3,557,587	2,993,383	84%

Notes:

- HCAI Patient Discharge Data based on data reported by facilities to HCAI (Hospital Inpatient - Characteristics by Facility (Pivot Profile) - Dataset - California Health and Human Services Open Data Portal)
- Differences in the methodologies between the HPD System and the comparison source may contribute to the variances displayed in the exhibit.

Exhibit 16 shows the number of individual physicians represented in the HPD System compared to external sources. For calendar year 2020, the HPD System contains services provided by over 83,000 individual physicians, which is comparable to an external source on physicians providing care in California.

Exhibit 16. Physicians Represented in the HPD System, 2020

MEASURE	NUMBER OF PHYSICIANS
Individual physicians with at least 10 services in the HPD System	83,015
Comparison Data:	
• Active California license and practice in California	123,941
• Provided patient care at least one hour per week	88,145
• Provided 20 or more hours of patient care per week	75,468

Notes:

- Individual physicians in the HPD System measured by counting the number of unique Rendering National Provider Identifiers (NPIs) with at least 11 services in 2020.
- Comparison data from J. Coffman, M. Fix, *The State of California's Physician Workforce*, Healthcare Center at UCSF, 2021.

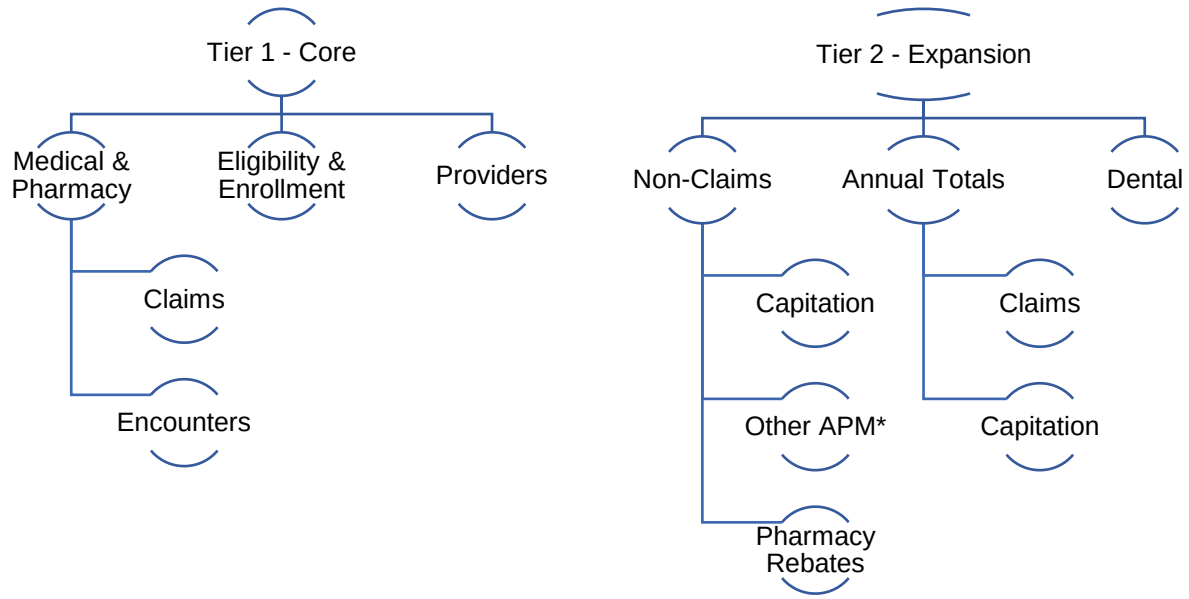
4. Opportunities to Expand the HPD

This section discusses potential ways to increase the types and amounts of data made available through the HPD Program. To maximize its utility and value for California policymakers, researchers, and others interested in improving California's healthcare system, HCAI intends for the HPD to be as comprehensive and complete as possible by increasing the quality, volume, and variety of data collected over time. The HPD, even in the first iteration, is a tremendous asset for anyone interested in better understanding—or improving—California's healthcare system. Representing 82% of Californians and 89% of all covered lives in the state, the HPD System is the largest APCD in the country. Opportunities to further expand the HPD are described below and categorized into two areas: additional data types (e.g., dental and non-claims payment data including capitation, other alternative payment arrangements, and pharmacy rebates), and additional submitters (e.g., additional voluntary self-funded plans, provider submitters, and supplier submitters).

Additional Data Types

For the initial implementation of the HPD System, the HPD Program focused on "core" files commonly available in administrative data and used in other state APCDs: medical and pharmacy claims and encounters, eligibility and enrollment records, and data about providers. Over the next several years, as required by HPD's enabling statute, HCAI plans to expand data collection to include dental data and non-claims payment data. Exhibit 17 displays the initial data collection effort as "Tier 1 - Core" and "Tier 2 - Expansion."

Exhibit 17. Tiers of HPD Data Collection Efforts



*APM = Alternative Payment Model

Add Dental Data

Dental coverage is typically offered as a standalone product alongside medical coverage and features a distinct set of market players and characteristics. The HPD Program will begin collecting data—in the standard APCD-CDL™ format for dental claims, eligibility, and provider files—from dental plans and insurers in 2024. As with medical and pharmacy data collected in Tier 1, dental data collection will begin with test files, proceed to submission of historical data back to June 2017, and then shift into routine monthly submissions on an ongoing basis.

Add Non-Claims Payment Data

The core data collected during the initial implementation of the HPD System includes claims and encounters from healthcare payers. These data capture payments made on a FFS basis from health plans and insurers to providers do but do not capture non-FFS payments, such as capitation, that play a major role in California's market.[iv] Millions of Californians across all payer types (commercial, Medi-Cal, and Medicare) are enrolled in managed care plans that use some type of non-claims payment for at least some of the services provided to their members. Collecting payment information for these services is essential to gaining a complete picture of the total cost of care in California. Total cost of care calculations will support analysis of variation by geographic region, payer, and product; total cost of care will also serve as the denominator for assessing the share of spending on primary care and behavioral health. Capitation and other non-claims payment data will also support a wide range of research questions such as the comparative effectiveness and cost of different models of care. Additionally, non-claims payment data is central to multiple OHCA use cases, including measuring adoption of alternative payment models that provide financial incentives for equitable high-quality and cost-efficient care.

Non-claims payment data that the HPD Program plans to begin collecting include:

- ***Capitation.*** These payments are population-based, usually per member per month, that may cover professional services, professional and facility services, or other services as negotiated by the plan and provider.
- ***Alternative payment arrangements other than capitation.*** These payments are contract-based (vs. member-based), vary by plan, and may be complex. Generally intended to shift payments from reimbursement of specific services toward value, examples include performance incentives and shared savings/risk arrangements.
- ***Pharmacy rebates.*** These payments are typically at the drug class level and reflect price concessions paid by a drug manufacturer to a pharmacy benefit manager or health plan.

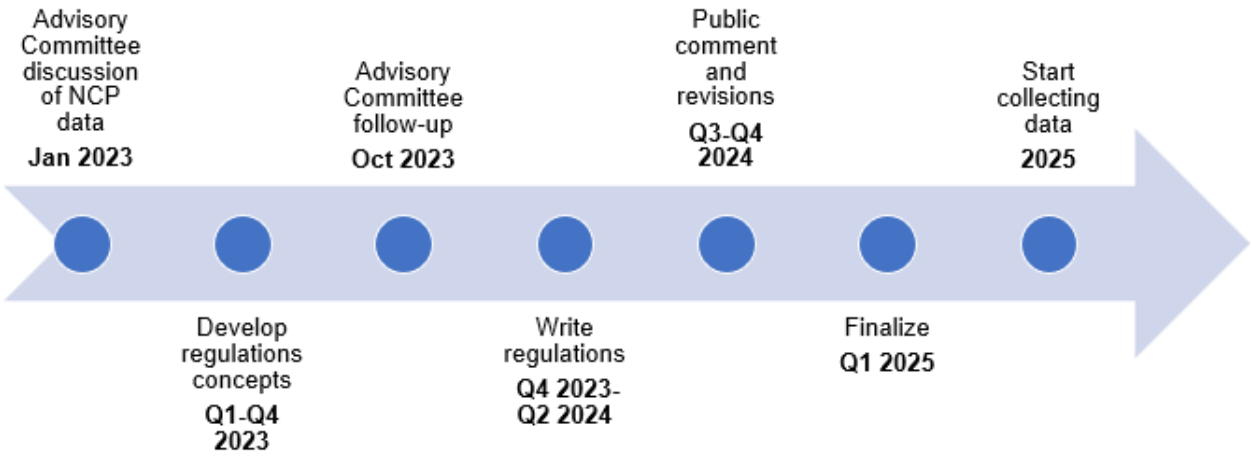
Unlike claims, encounters, and eligibility data, a national standard for collection of non-claims payment data has not yet been developed, but HCAI is working closely with the National Association of Health Data Organizations (NAHDO) and the APCD Council to develop a data layout that could be incorporated into the APCD-CDL™.

[iv] Subsections in Section 5 of this Report contain additional information about data on claims and encounters: Background on Administrative Data in Health Care and Impacts to Data Quality, and Encounter Data Completeness in the HPD.

HCAI continues to discuss specifics and approaches to collecting non-claims payment data with the HPD Advisory Committee, NAHDO, and health plans and insurers, and intends to coordinate the needs of both the HPD Program and OHCA.

A projected timeline for those efforts is shown in Exhibit 18 below.

Exhibit 18. Timeline For Non-Claims Payment (NCP) Data Collection



Additional Submitters

Add More Voluntary ERISA Self-Funded Plan Submitters

The HPD System includes detailed healthcare service, payment, and enrollment information for over 31 million individuals per year. Exhibit 19 provides details on the population in and out of the HPD System for the most recent complete year in the HPD System. For calendar year 2021, the HPD System includes an average of 32.3 million members per month, or 82.4% of California's population.

The following groups, accounting for approximately 6.9 million people, are not included in the current iteration of the HPD System:

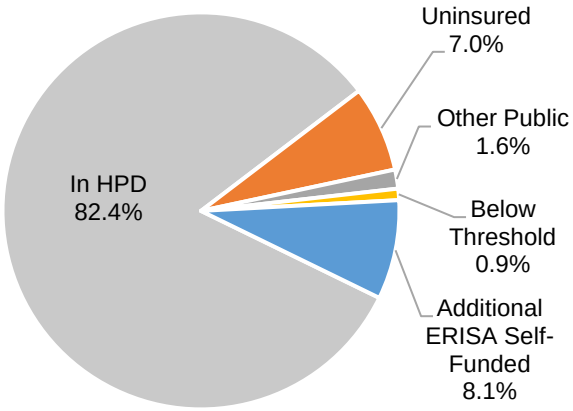
- Uninsured
- Other public programs (e.g., federal employee and military)
- ERISA self-funded plans exempt from state mandate (e.g., private self-funded employers and Taft-Hartley trusts)
- California licensed plans and insurers that fall below the threshold for mandatory submission to the HPD Program (less than 40,000 non-Medi-Cal covered lives)

Exhibit 19 displays the estimated proportion of Californians in and out of the HPD System for calendar year 2021. The HPD System includes healthcare service and eligibility data for 82.4%

of all Californians. The biggest category not included—ERISA self-funded plans not subject to mandatory submission—accounts for approximately 3.2 million Californians.

Exhibit 19. Percent of Californians Represented in the HPD System, 2021 (Estimated)

POPULATION GROUP	NUMBER	%
Included in the HPD	32,376,087	82.4%
Not Included in the HPD		
Uninsured	2,749,344	7.0%
Other Public (e.g., Military, Federal Employees, Indian Health Service)	609,000	1.6%
Below Threshold	365,428	0.9%
Additional ERISA Self-Funded	3,176,484	8.1%
Total Californians	**39,276,343**	**100%**



Notes:

- Number of uninsured from US Census Bureau.
- Number in Other Public from California Health Benefits Review Program, *Estimates of Sources of Health Insurance in California, 2021*.
- Number below threshold based on HCAI analysis of covered lives reported in the California Health Care Foundation, California Health Insurers, Enrollment, 2023 Edition and HPD Program mandatory reporting thresholds. Includes regulated health plans and insurers only. A health plan, health insurer, or public self-insured plan that has fewer than 40,000 California members is not required to submit data to the HPD Program.
- Number in ERISA Self-Funded estimated from HCAI analysis and represents *additional* ERISA covered lives not already included in the HPD System. Derived by subtracting other categories from Total Californians. Note this may also include a small number of covered lives in public self-funded plans.
- Total Californians from US Census Bureau.

Background on ERISA

In March 2016, the United States Supreme Court ruled that states cannot require self-insured or self-funded employer plans regulated under the Employee Retirement Income and Security Act of 1974 (ERISA) to submit data to a state APCD. The decision, *Gobeille v. Liberty Mutual*, resulted from a lawsuit by a self-insured employer that challenged Vermont's right to require the employer's Third Party Administrator (TPA) to submit claims data to the state APCD. The Supreme Court found that ERISA preempted Vermont's ability to compel the submission of claims data for self-funded employers.[17] As a result, state APCDs have struggled to collect data from entities noted in the decision.

The Supreme Court ruling applies to entities that are subject to ERISA and are self-funded. Two groups are worth noting:

- **Private self-funded employers**. Many large firms opt to bear the financial risk associated with providing health benefits for their employees, rather than paying an insurance company to do so.
- **Taft-Hartley trusts.** Collectively bargained and administered by an equal number of management and union representatives, Taft-Hartley trusts are generally multi-employer

arrangements and are often used to cover workers in project-based or seasonal jobs such as construction.[18]

Private self-funded employers and Taft-Hartley trusts cannot be compelled to submit data, but they can do so voluntarily. The table below summarizes the key distinctions in determining whether submission of data on self-funded lives to the HPD Program is mandatory or voluntary. Public entities that self-fund health benefits are not subject to ERISA, so they are required to submit to the HPD if they have over 40,000 self-insured covered lives in their self-funded plans. Examples of public self-funded plans registered as submitters to the HPD include plan options from CalPERS.

Exhibit 20. ERISA and Plan Size Determine Whether Submission of Data on Self-Funded Plans to HPD is Mandatory

ERISA STATUS	<40K SELF-FUNDED COVERED LIVES	>40K SELF-FUNDED COVERED LIVES
Non-ERISA (subject to California state law mandating data submission to HPD)	**Voluntary**	**Mandatory** Example: CalPERS
ERISA (regulated by federal Department of Labor, exempt from California state law mandating data submission to HPD)	**Voluntary** Example: employers and labor trusts	

Assessment of Self-Funded Data in the HPD

Self-funded plans generally contract with a vendor, often a carrier, to provide "administrative services only" (ASO) functions such as claims processing but not insurance. Self-funded lives can be measured in the HPD System using the APCD-CDL™ element Coverage Type which includes two values for submitters to identify members in self-funded arrangements:

- **ASW** (ASO with Stop-Loss): self-funded plans administered by a TPA, where the employer has purchased stop-loss, or group excess insurance coverage.
- **ASO** (ASO without Stop-Loss): self-funded plans administered by a TPA, where the employer has not purchased stop-loss, or group excess insurance.

Though there is no definitive source of the total number of self-funded covered lives in California, the best plan-specific estimates come from DMHC and CDI reports. These reports include total self-funded lives and are not available separately for ERISA vs. non-ERISA. Similarly, the APCD-CDL™ does not include an element to differentiate ERISA vs. non-ERISA in the data that submitters send to the HPD Program. State regulators have limited oversight of ERISA self-funded plans, and carriers may not have ready access to the ERISA status in the systems used to prepare data for APCDs. Accordingly, state APCD programs struggle to definitively report on ERISA self-funded separately from all self-funded enrollment. For the purpose of this report, HCAI used methods of estimation to gauge the degree of ERISA inclusion in the database. State-wide estimates were made using the United States Department of Labor reports on private vs. public self-funded lives in California[19] which show that 78% of

self-funded lives are private, and a combination of Coverage Type, Plan Name, and Group Number/Names to estimate the ERISA lives in the HPD System.

Exhibit 21 shows the estimated number of self-funded lives state-wide and in the HPD System. Using the estimation methods described above, approximately 4.5 million out of 5.8 million self-funded lives state-wide in 2022 are ERISA. The HPD System includes approximately 300,000 ERISA lives and 1.1 million total self-funded lives in 2022. The rest are from public self-funded plans such as CalPERS, the University of California, Self-Insured Schools of California, and California's Valued Trust. Overall, for 2022 the HPD System includes approximately 7% of the ERISA lives and 20% of all the self-funded lives in the state. Please note again, however, that these are estimates, and based on the prior analysis summarized in Exhibit 19 that indicates there are 3.2 million ERISA self-funded lives not in the HPD, the 1.1 million lives with a self-funded Coverage Type displayed in Exhibit 21 may be underreported.

Exhibit 21. Estimate of Self-Funded Covered Lives, in Millions

CATEGORY	2018	2019	2020	2021	2022
State-Wide					
ERISA	4.3	4.4	4.4	4.3	4.5
Non-ERISA	1.2	1.2	1.2	1.2	1.2
Total, State-Wide	**5.5**	**5.6**	**5.6**	**5.5**	**5.8**
In HPD System					
ERISA	0.2	0.2	0.2	0.2	0.3
Percent of State-Wide ERISA	4%	4%	5%	5%	7%
Non-ERISA	0.8	0.8	0.8	0.8	0.8
Total, HPD System	**0.9**	**1.0**	**1.0**	**1.0**	**1.1**
Percent of State-Wide Self-Funded	17%	17%	18%	19%	20%

Notes:

- State-Wide totals based on DMHC and CDI data from California Health Care Foundation, California Health Insurers, Enrollment – 2023 Edition.
- State-Wide ERISA and Non-ERISA estimated based on percent of Private Sector and Public Sector self-insured lives reported by the Department of Labor, 2021 Health Insurance Coverage Bulletin.
- HPD System figures based on HCAI estimate, using eligibility records in December of each year with Coverage Type = ASW or ASO and an analysis of Plan and Group names.

Preliminary analysis of the self-funded lives in the HPD indicates that voluntary participation of ERISA plans is low and that most ERISA self-funded lives are not included in the HPD.

HCAI's Approach to Encouraging Voluntary Submission of Self-Funded Data

To encourage voluntary data submission, HCAI engaged in a series of meetings with self-funded employers and other purchasers to learn about their perceived value of HPD data and motivating reasons for voluntary submission. In 2021, during the planning phase of the HPD Program, HCAI met with employer coalitions, including the Purchaser Business Group on Health, Silicon Valley Employers Forum, Catalyst for Payment Reform, and the Maine Health Purchaser Alliance. HCAI also gathered insightful information from several state APCDs and experts, including Colorado's Center for Improving Value in Health Care (CIVHC), staff from Utah's APCD, the APCD Council, and the Integrated Healthcare Association.

In addition, HCAI collaborated with the California Health Care Coalition (CHCC), a membership organization of public- and private-sector employers, unions, and health and welfare trust funds. On behalf of HCAI, CHCC conducted a survey of its membership soliciting high-priority use cases for HPD data. This collaboration resulted in HCAI creating resources that support employers and other purchasers interested in voluntary submission of data to the HPD Program, such as information about the benefits of opting-in and guidance on how to instruct their ASOs to voluntarily submit data to the HPD Program. The resources, including a FAQ and a request form, were published on the HCAI website in June 2022. To date, HCAI has not received any requests for support with voluntary submission of data to the HPD Program.

HCAI also met with the six major California carriers that provide ASO services to better understand their ASO lines of business, learn about how the national carriers work with APCDs in other states, and enlist their assistance in obtaining authorization from self-funded employers and other purchasers to submit data to the HPD Program. HCAI encouraged the carriers to proactively inform their clients about the opportunity to submit data to the HPD Program and to implement processes for employers to easily authorize submission of their data. The carriers vary in their approach. Some include an "opt in" authorization process in their client onboarding or annual renewal processes, a more active approach. Others rely on requests from their clients to trigger the authorization and opt-in process, which is more passive. One carrier cited concerns about sharing data from ERISA self-funded plans, even if requested by their clients. Costs may also be a barrier to submitting data for ERISA plans—there have been anecdotal accounts nationally of TPAs charging a fee to self-insured clients for submitting data to APCDs.

There are multiple reasons California ERISA self-funded employers and other purchasers may choose to contribute their data to the HPD. First, many health benefit programs are increasingly concerned about the continued escalation of health care costs and are interested in more effective use of data and transparency to drive the development of solutions. The more comprehensive the HPD is in reflecting the entire population of California, the more accurate the analytic findings will be and the greater the likelihood of achieving both purchaser goals and HPD program goals—including transparency, informed policy decisions, improved health care, and equitable access to affordable and high-quality care.

Second, HPD features more robust data than is available from a data warehouse or health plan, and can support employers and other purchasers in achieving organization-specific cost,

quality, and equity goals. A partial list of high-value use cases that California purchasers have raised in discussions about HPD include:

- Benchmarking information for comparative evaluation of networks, including total cost of care data and comparative information about high-volume procedures
- Data on provision of low-value care
- Assigning a cost to capitated encounters
- Detailed cost data on prescription drugs
- Site of service analysis, incorporating quality outcomes
- Integrating data on race/ethnicity and language to illuminate disparities, increase health equity

With the initial HPD Program implementation complete and the first public reports released in June 2023, HCAI has an opportunity to perform additional targeted outreach aimed at large employers and other purchasers. Making a compelling case for voluntary submission of data was challenging while the HPD Program was still in development; HCAI can more easily demonstrate the analytic power and potential of the HPD System now that the initial system has been developed and implemented. HCAI is considering next steps such as:

- Follow up with ASO plans to share HPD progress to date and to encourage increased data submission.
- Meet with partners in the purchaser community to share data on inclusion and representation of this segment of the healthcare marketplace and to cultivate increasing participation.
- Share progress to date with Taft Hartley plans and large employers to increase awareness of the HPD and its potential value to them; the value will increase with additional participation.
- Outreach to other potential interested parties such as benefits consultants who may provide support to large employers with their benefit plans.

Additionally, state policymakers could take steps that place requirements on TPAs that increase likelihood of submission of ERISA self-funded data to the HPD. For example, Utah has regulations that obligate insurers in the state to provide an opt-in form to their self-funded employer clients, and carriers must submit data to the APCD for any employer clients that opt-in. Utah carriers also must annually report the number of employers that opted in and out, the name and contact information for employers provided the opt-in form, and certify that they made reasonable efforts to provide the form to all known required employers.[20] Additionally, several states are considering polices that limit the fees ASOs are able to charge to submit data to the APCD.

Add Providers and Suppliers as New Submitters

The HPD Program's data collection efforts to date have focused on the payers and plans in California: DHCS for Medi-Cal, CMS for Medicare FFS, and regulated health plans and insurers for commercial and Medicare Advantage data, consistent with the recommendations the HPD Review Committee made in 2020. There are several reasons for this approach, many of them practical. Payers and plans have data about healthcare services provided, have experience

sharing the data with others, and represent a considerably more manageable number of entities and sources from which to coordinate. For those reasons and others, states use payers and plans as the primary source of data for their APCDs. Additionally, HCAI already collects certain information from providers and maintains several databases of hospital data (See ED Visits and Inpatient Discharges subsection above).

California's statute allows the HPD Program to collect data from other healthcare entities, including providers and suppliers, as defined below:

- **Provider** means a hospital or a clinic (HSC, § 127673, subd. [d][4])
- **Supplier** means a physician and surgeon or other healthcare practitioner, or an entity that furnishes health care services other than a provider, that has an independent scope of practice and submits claims electronically (HSC, § 127673, subd. [d][5])

As described in previous sections of this report, the HPD Program already collects provider and supplier data indirectly via payers and plans; it is unlikely that the number of healthcare services captured in the HPD System would increase significantly from a new data collection effort focused on providers and suppliers. Adding provider and supplier organizations to the list of required submitters, assuming they were to submit the same types of data files currently required of plans and payers, would also add considerable cost and complexity to the operation of the HPD Program, including:

- **Exponential increase in data feeds**. The number of submitters would increase from approximately 50 organizations to potentially tens of thousands, including more than:
  - 200 risk-bearing organizations[21]
  - 500 hospitals[22]
  - 16,000 physician group practices[23]
  - 100,000 individual physicians[24]
  - 1000s of home health agencies, clinics, and other healthcare entities
- **Data duplication implications**. Most of the data collected from providers and suppliers would be duplicative of the information already collected from plans and payers, requiring a complex ongoing evaluation and de-duplication process to ensure that services and payments are not overstated.
- **Increased costs and reporting burdens for providers and suppliers**. Unlike plans and payers, many providers and suppliers do not have the resources or experience to share the types of data required for the HPD Program. HCAI previously estimated one-time costs to each submitter of $307,000 and annual costs of $133,000.[25] Those estimates were for large health plans experienced with data sharing; costs for providers and suppliers without such experience may well be different and could represent a more significant impact for a smaller organization.
- **Increased costs for the HPD Program**. An exponential increase in the number of submitters would require significant new resources to manage the data collection, evaluation, and integration processes and add considerably to the estimated $22 million annual expenditures to operate the current HPD Program.[2]

An alternative approach to data collection from providers and suppliers could be considered, focused on certain types of payment data from risk bearing organizations (RBOs, see inset). RBOs are physician organizations that enter into contracts with health plans to provide services to members on a capitated basis, and contract with additional providers (often called "downstream" providers) for a subset of those services. The payments made by RBOs to their downstream contracting partners take a variety of forms, such as sub-capitation, case rate payment, and FFS. While HCAI anticipates beginning collection of non-claims payment data for the HPD Program, such as capitation from health plans to RBOs and other provider organizations beginning in the third quarter of 2025 (see the previous subsection, Add Non-Claims Payment Data), the payments made by the RBOs to their downstream contracting partners are not currently captured by the HPD Program and will not be included in the non-claims payment submission from payers and health plans.

Capturing payment data from RBOs to downstream providers would help complete the total cost of care picture, supporting comparison of the performance of different models of care. It could also support drill-down analysis on categories of service, such as primary care or behavioral health services. Other potential uses of the RBO data include:

- Benchmarking, e.g., comparative data on payment for various specialty services.
- Validate encounter data completeness: data is likely most complete at the RBO level; there may be some loss given that the data goes through multiple handoffs (e.g., clearinghouses) between the RBO and HPD.
- Enable comparison of RBO cost vs. health plan cost (i.e., capitation paid to RBO) for the same services.
- Spend on primary care—understanding and quantifying primary care services provided by RBOs that are not billable, e.g., pharmacist-run programs that reduce costs and complications but won't be captured in the primary care spend because the pharmacist is not a billable provider.

What is a Risk Bearing Organization?

A risk bearing organization (RBO) is either a professional medical corporation, other form of corporation controlled by physicians and surgeons, a medical partnership, a medical foundation exempt from licensure pursuant to subdivision (l) of Section 1206 of the Health and Safety Code, or another lawfully organized group of physicians that delivers, furnishes, or otherwise arranges for or provides healthcare services. An RBO does not include an individual or a healthcare service plan. An RBO does all of the following:

- Contracts directly with a healthcare service plan or arranges for healthcare services for the healthcare service plan's enrollees.
- Receives compensation for those services on any capitated or fixed periodic payment basis.
- Is responsible for the processing and payment of claims made by providers for services rendered by those providers on behalf of a healthcare service plan when those services are covered under the capitation or fixed periodic payment made by the plan to the risk bearing organization.

As of June 2023, there are 212 RBOs registered in California with collective enrollment of 9.5 million.

———

Sources:

- California Department of Managed Health Care, Risk Bearing Organization (RBO) Frequently Asked Questions
- California Department of Managed Health Care, Provider Solvency Quarterly Update

In summer 2023, HCAI conducted a series of exploratory discussions with RBOs to increase understanding of the RBO data landscape and feasibility of a potential future effort. The focus of the discussions was the flow of data between plans, physician organizations, management services organizations (MSOs), and clearinghouses to inform thinking about what might be possible.

HCAI found that collecting data from RBOs would be a significant effort for the HPD Program and a burden for RBOs. HCAI may consider a pilot project with a small number of RBOs to evaluate this further.

There are 212 RBOs registered with DMHC,[21] so the number of registered submitters to the HPD Program would increase five-fold. An annual (vs. monthly) data collection effort focused on claims and non-claims payment data would help reduce the cost and burden. Other challenges would also need to be overcome, including potential proprietary concerns from RBOs regarding provider networks and payment rates.

5. Data Quality

This section presents an initial high-level evaluation of the quality of the data collected and stored in the HPD System. There are several attributes of data quality for an APCD, including data completeness, accuracy, and timeliness. Other portions of this Report, in particular the comparisons of visits, discharges, and prescriptions to external benchmarks presented in previous sections, provide encouraging indications about data completeness and accuracy. The sections on submitter data that show consistency in record and member counts over time are also good initial measures of completeness and timeliness. The sections below discuss additional aspects of data quality, including how administrative data processes impact quality, how data quality processes are integrated into the HPD Program's data collection and feedback processes, and how general measures of data element and encounter data completeness in the HPD compare to other sources.

Background on Administrative Data in Health Care and Impacts to Data Quality

The HPD Program, like all APCDs, collects and aggregates administrative data from healthcare payers including CMS, Medi-Cal (DHCS), and commercial health plans and insurers. By its nature, administrative data is not intended for use by researchers in analytic databases such as an APCD, but it does provide rich analytic value and represents the most accessible source for the detailed healthcare services and payments provided in a healthcare system. The quality of the data available in an APCD is dependent on the quality of the underlying administrative data and processes used by payers and providers to process healthcare claims and on the completeness of the encounter data reported and captured in plan and payer systems.

The primary function of administrative data in health care is program operations. Providers submit claims for payment for services provided to patients. The insurer or administrator checks to make sure the patient is eligible under their health plan to receive the services provided, determines the fee the provider is owed, deducts the patient's share of the payment, and issues payment to the provider. This process is called claims adjudication.

Under capitation, providers do not submit claims for individual services; they are instead expected to submit encounters to document provision of services. Encounters and claims both

What's on a Member Eligibility Record?

In carrying out their healthcare operations, plans, insurers, and administrators use demographic information about members eligible for services under their plan, including:

- Name
- Sex
- Date of birth
- Address
- Type of plan/coverage
- Numeric identifiers (e.g., plan ID, SSN when available)
- Dates of eligibility

What's on a Health Care Claim or Encounter?

When patients see their healthcare provider, the provider will often submit a *claim* for to the patient's health plan to request payment for the healthcare services provided. A similar type of record, an *encounter*, captures information about healthcare services received when the provider does not require direct payment for those services. Claims and encounters include:

- Patient and Provider identifiers
- Dates of service
- Location where service was provided
- Diagnosis codes
- Revenue codes
- Pharmacy codes
- Patient cost sharing
- Payment, if applicable

include information about the service, diagnosis, patient, and provider information; the primary difference is that claims are used for billing and payment while encounters are not.

Data quality for certain types of administrative data, notably those required to successfully adjudicate a claim or encounter, is quite high—for example, elements that relate to the identification of the patient, service date, payment date, procedure codes, and identification of the billing provider. The use of standard claim forms, first on paper and now almost exclusively electronic, also helps improve data quality for these core elements. The required use of standard electronic healthcare transactions, formalized by the 1996 Health Insurance Portability and Accountability Act, specify the format, valid values, and usage of the key elements required in transactions between plans and providers, many of which are also critical to the utility of a research database like an APCD.

In contrast, data elements that are not required or used in administrative processes may not be consistently collected, stored, and accessible to healthcare plans and insurers when creating data for an APCD. Examples of these elements include diagnosis codes other than the primary diagnosis, race, ethnicity, and language. For these elements, there is often more variation in the completeness and quality of the data stored in plan and payer underlying systems.

How Will the HPD Program Ensure Data Privacy, Confidentiality, and Security?

Chapter 8.5 (commencing with Section 127671) of Part 2 of Division 107 of the Health and Safety Code establishes strong privacy and security standards for the HPD System.

Access to personally identified data collected by the HPD is only authorized for eligible uses and after appropriate approvals have been granted. Only eligible and authorized entities, such as university researchers, may obtain appropriate access to individual-level data for the purposes of data linkage and longitudinal research study. The HPD data may not be used for determinations regarding individual patient care or treatment nor for any individual eligibility or coverage decisions. The HPD data are exempt from the California Public Records Act.

See also Section 6, Public Reporting and Data Release.

It's important to note that although the HPD collects very detailed information on patients and services, including name, date of birth, dates of service, and diagnosis codes, there are strict processes in place to safeguard this information (see inset, How Will the HPD Program Ensure Data Privacy, Confidentiality, and Security?).

The sections below focus on an assessment of field-level completeness, although other sections of this Report, in particular the comparisons of visits, discharges, and prescriptions to external benchmarks, provide encouraging indications about data completeness and accuracy. The following sections discuss the data collection and validation processes used by the HPD Program, the levels of data element completeness and variability by submitter, race and ethnicity data in the HPD, and encounter data completeness.

Data Collection and Validation Processes in the HPD

The HPD Program, in concert with their APCD Platform partner, works with HPD Program submitters to help ensure the best available data makes it into the HPD System. These activities include the following key components:

- Frequent and proactive communication with submitters
- Guidelines and requirements for data submissions, formalized in the HPD Program Data Submission Guide[26]
- Timely validation checks and communication to submitters about data that does not pass initial quality checks
- Evaluation of accepted data for variation in quality over time
- Acceptance and incorporation of replacement data, when appropriate, to address data quality concerns

HCAI has been in regular contact with submitters since the planning phases of the HPD Program. The Submitter Group, comprised of representatives from California's health plans and insurers, began quarterly meetings in 2021. The Submitter Group provides a forum for submitters to receive up-to-date information on submission requirements, troubleshoot data submissions, and address technical issues.

The HPD Program Data Submission Guide identifies the standard formats to use for each file type [All Payer Claims Database Common Data Layout (APCD-CDL™)] and further specifies criteria that each file must pass in order to be accepted for initial processing. Data elements designated as "required" must be populated for a record to be accepted. Data elements designated as "situational" must be populated under specific circumstances. For example, the claims file data element "Admission Date" is designated as "Situational" and is required when the claim or encounter is "Inpatient." Unless a variance has been registered and accepted for a specific element, failure to comply with the requirements will result in the rejection of the submitted file. Even when a data element is not designated as required or situational, population of these fields is still expected if the submitter has that data available.

Any file that fails initial validations is not accepted and must be corrected and resubmitted. Further edits are applied for specific data elements that use standard industry values [e.g., National Drug Codes (NDC), Healthcare Common Procedure Coding System (HCPCS), and the National Provider Identifier (NPI)] and where standard values are required from the APCD CDL™ data specifications.

As described in previous sections of this Report, California's health plans and insurers have worked closely with HCAI to submit the required data to the HPD Program, and HCAI is in regular communication with each of those submitters. As of the writing of this Report, all mandatory plans and insurers have successfully complied with the data submission requirements of the HPD Program. Over time, the types of data quality checks and the thresholds for acceptable data may need to be adjusted, and HCAI will work closely with submitters to communicate and implement any such changes. The law allows for HCAI to work with DMHC and CDI, as appropriate, to take action necessary to bring the submitter into compliance with the HPD Program's data requirements, as necessary (HSC, § 127674.1).

After all validations have been performed, an evaluation is made of data completeness for each field: the number of valid values is expressed as a percentage of the total expected values in

the file. This completeness percentage is then compared to HCAI reporting thresholds for each field. If the completeness percentage falls below the HCAI reporting threshold, the file is not accepted. Submitters must either: a) correct and resubmit, or b) request a data variance for approval by HCAI.

Depending on tolerances for data completeness, the analytic dataset will include some instances of missing or invalid data. Additional filters are applied before readying the data for analysis and access by users of the HPD System. For example, enrollment and service records are excluded for individuals living outside of California, product types not yet accepted (e.g., dental plans, Medicare supplemental plans, student health plans, etc.), and dates of service outside of the reporting period. All of these exclusions and filters are clearly indicated in documentation made available to users of the data.

The HPD Program's data validation efforts benefit from the mature data quality processes used upstream from the HPD System. For example, in accordance with federal regulations, DHCS contractually requires that Medi-Cal Managed Care Plans submit to DHCS complete, accurate, and timely encounter data for services provided to enrolled beneficiaries,[27] and since 2015 DHCS has only accepted national standard file formats and coding schemes for managed care encounter data submissions. Health plans and insurers provide data to others within the state (e.g., Integrated Healthcare Association (IHA), Covered California, and CalPERS) that apply their own data quality and validation checks. Use of the data in analytic databases in each of those organizations also likely serves to, in the long term, increase quality and completeness of the data state-wide. Collectively, these submission, validation, and use efforts improve the quality of data throughout California's healthcare system.

Data Element Completeness in the HPD

This section provides information about the completeness of specific data elements stored in the HPD System—for example, the proportion of records submitted with a valid value. As described above, the HPD Program Data Submission Guide identifies required data elements, or elements that must be provided for every record unless an exception has been granted to the submitter. Required elements include elements that are typically available from submitters and commonly used by researchers in APCDs. Examples of required elements for eligibility, medical, and claims files are shown in Exhibit 22.

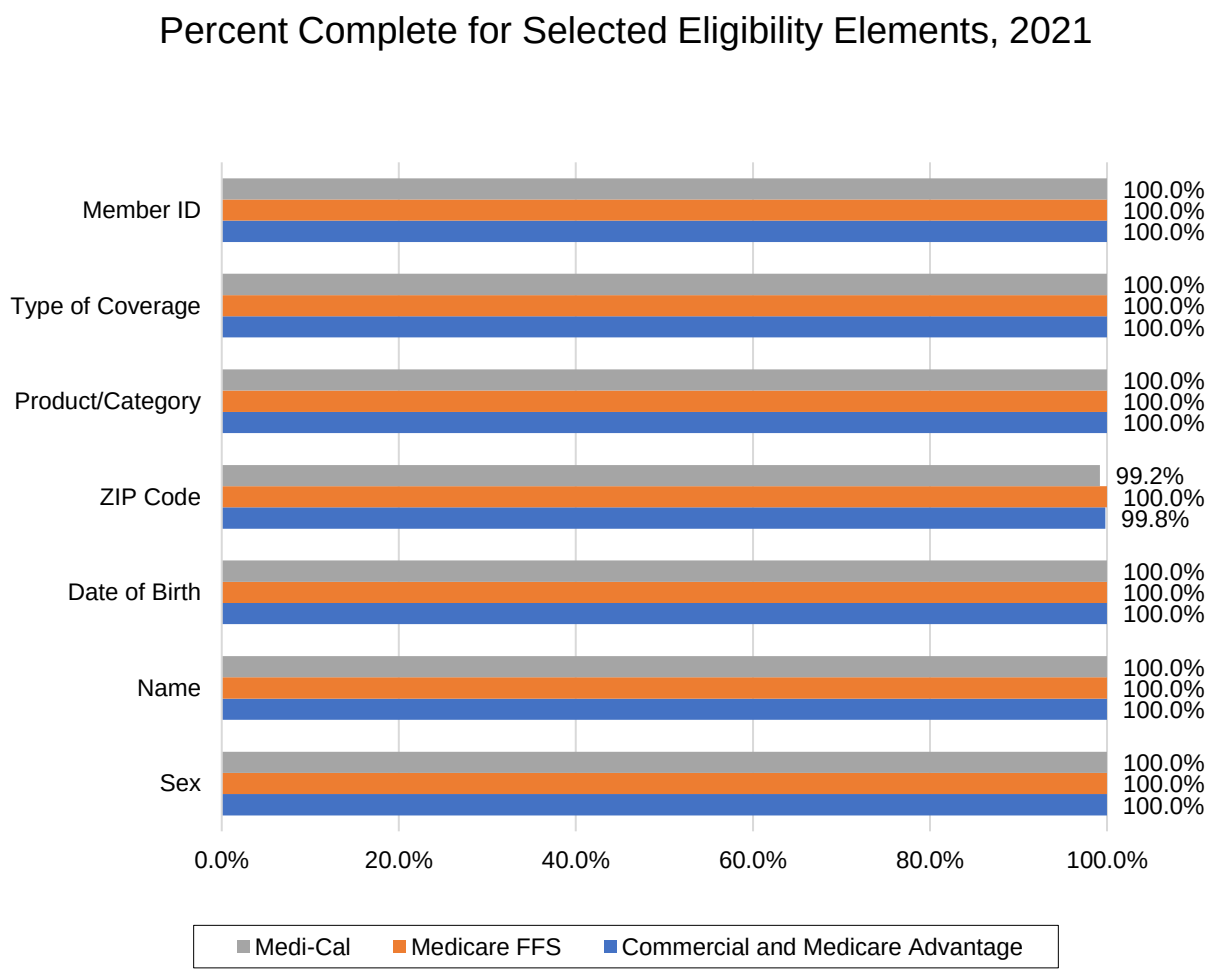
Exhibit 22. Examples of Required Elements by File Type

MEMBER ELIGIBILITY FILE	MEDICAL CLAIMS FILE	PHARMACY CLAIMS FILE
Member Name	Principal Diagnosis	Date Prescription Filled
Member Date of Birth	Procedure Code	Drug Code
Member Sex	Revenue Code (for institutional claims)	Quantity Dispensed
Member ZIP Code	Dates of Service	Days Supply

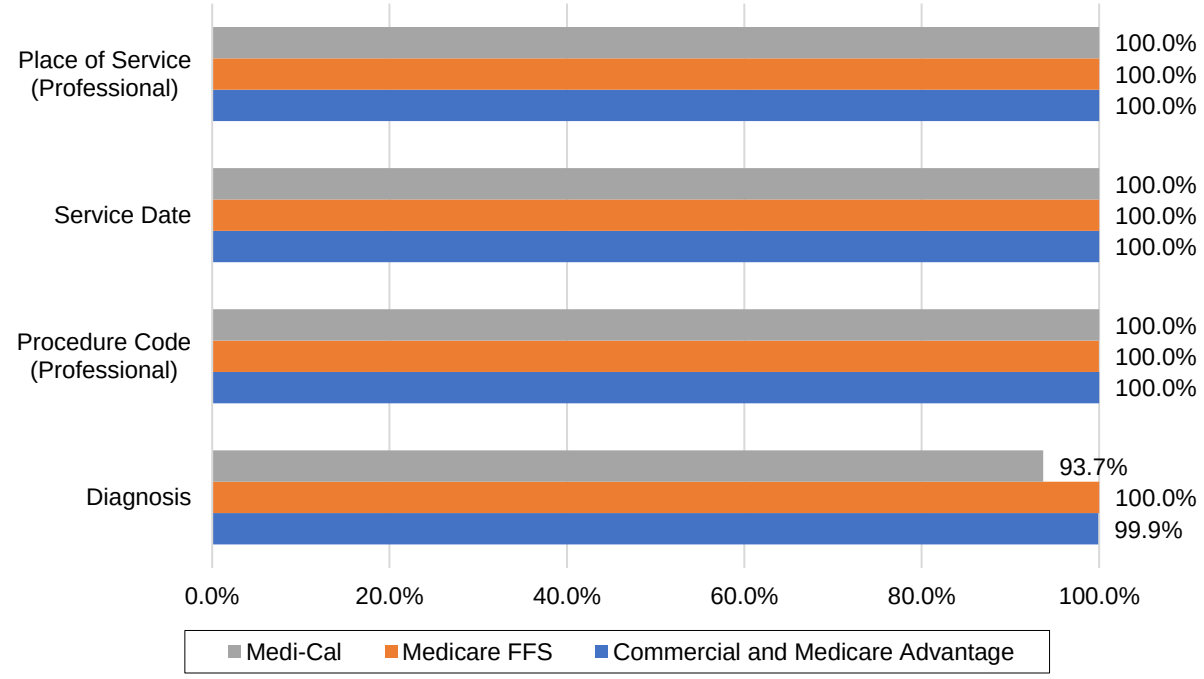
MEMBER ELIGIBILITY FILE	MEDICAL CLAIMS FILE	PHARMACY CLAIMS FILE
Member ID	Service Units	Prescribing Physician ID
Insurance Product/Category	Allowed Amount/ Fee-For-Service Equivalent	Pharmacy ID
Type of Coverage (i.e., Medical, Pharmacy, Behavioral Health)	Rendering Provider ID	Drug Unit of Measure
Data Submitter Code	Billing Provider ID	Pharmacy ZIP Code
Payer Code		

The graphs in Exhibit 23 below show the percent complete rates, by submitter type, for selected required fields in the eligibility, medical claims, and pharmacy claims files. As expected, completion rates are high for these key fields in the HPD System.

Exhibit 23. Percent Complete for Selected Eligibility, Medical, and Pharmacy Elements



Percent Complete for Selected Medical Elements, 2021



Notes:

- Principal Procedure Code (Facility) not displayed because they are only required when a significant procedure occurred during the hospitalization.

Percent Complete for Selected Pharmacy Elements, 2021

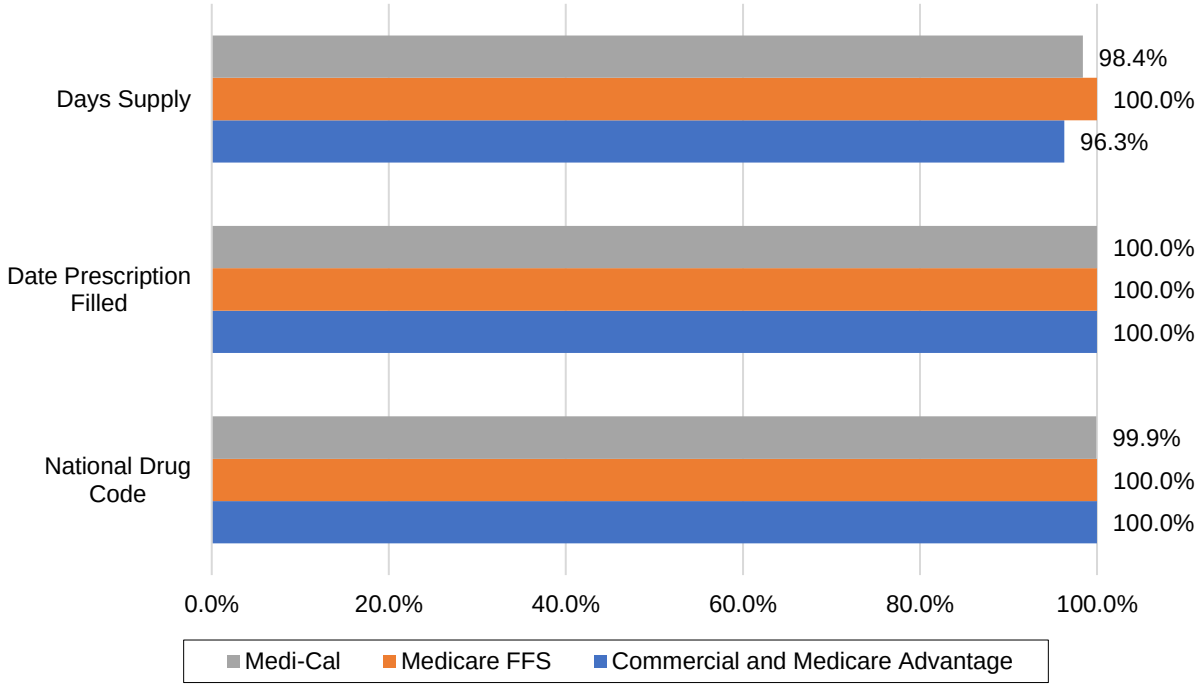
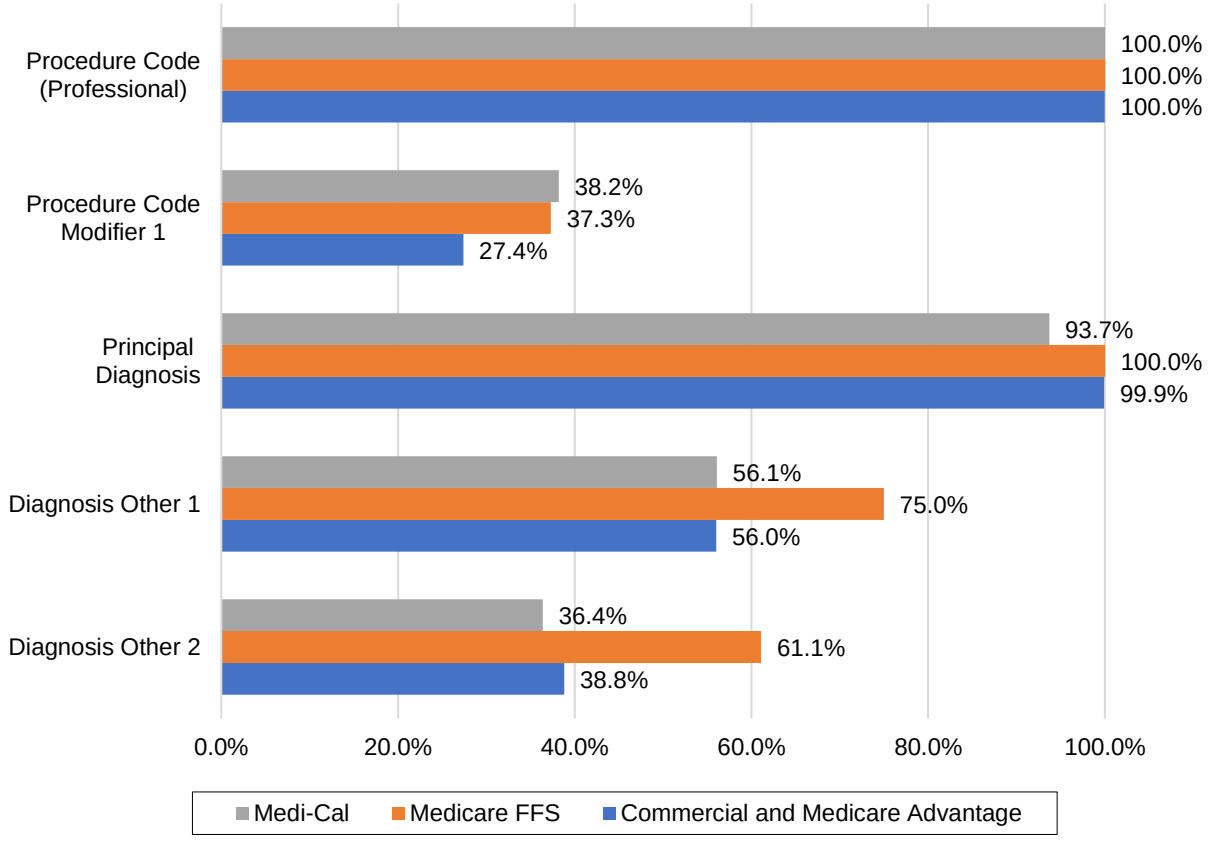


Exhibit 24 shows completeness levels for diagnosis and procedure codes. Procedure Code and Principal Diagnosis code are required elements and are present on virtually all records for all submitters. Procedure Code Modifier 1, Diagnosis Code Other 1, and Diagnosis Code Other 2 are situational elements that can also be important for analyses of healthcare conditions. Procedure Code Modifiers are used to provide additional information about the service and are only required for certain types of services. For example, a modifier may provide details not included in the code descriptor, such as the anatomic location of the procedure.[28] Additional diagnosis codes can provide important information about the patient's comorbidities that are not the primary reason for the service. Procedure code modifiers and additional diagnosis codes enhance the clinical value of the information in analytic databases like an APCD. The completion rates for these situational fields in the HPD System are generally similar across submitters and comparable to data in other APCDs.

Exhibit 24. Percent Complete for Selected Medical Elements: Procedure Code Modifier and Additional Diagnosis Codes (2021)



Data elements typically used by APCDs, including basic demographic information about the patient and the types and location of services, are consistently provided by submitters and support a wide variety of standard analyses.

Race, Ethnicity, and Language

Exhibits 25 and 26 provide completeness data for race, ethnicity, and language. In the context of growing recognition of disparities in health outcomes based on social drivers of health, data on these and other elements is increasingly important in understanding variation and addressing inequity. Yet, the data are often incomplete and non-standard, limiting their usability. Authors of a 2023 National Committee for Quality Assurance (NCQA) project, *Current Health Plan Approaches to Race and Ethnicity Data Collection and Recommendations for Future Improvements*,[29] concluded that:

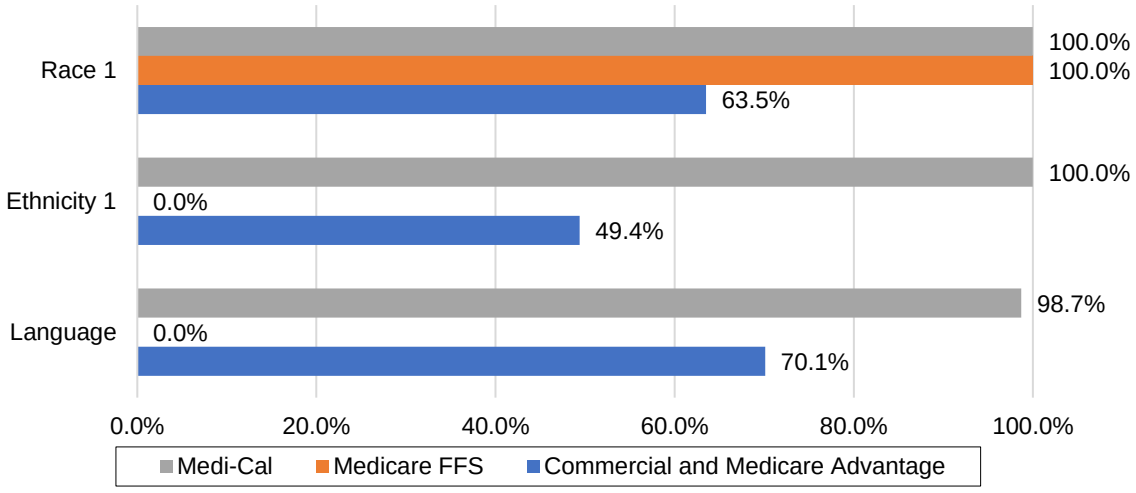
> *. . . collecting data remains difficult; there is a dearth of standards for how health plans should collect these data and a lack of standards for recording and reporting them. Currently, health plans collect data on race and ethnicity in a variety of ways: from government or private payers, through interactions with plan members, from patients' clinical records, and by attributing race and ethnicity based on name and place of residence.*[30]

The Medicaid and CHIP Payment and Access Commission reached similar conclusions about the quality of Medicaid race and ethnicity data in their 2023 report, *Medicaid Race and Ethnicity Data Collection and Reporting: Recommendations for Improvement*.[31] The U.S. Department of Health and Human Services Office of Inspector General documented concerns about Medicare's race and ethnicity data in a 2022 report.[32]

State APCDs struggle to collect useful race and ethnicity data from payers. In a survey of five state APCDs, in 2017 just 13%-44% of records contained a valid race value, with an average of 28%.[33] CIVHC, Colorado's APCD, reports that 84% of data from commercial payers has an unknown race value.[34]

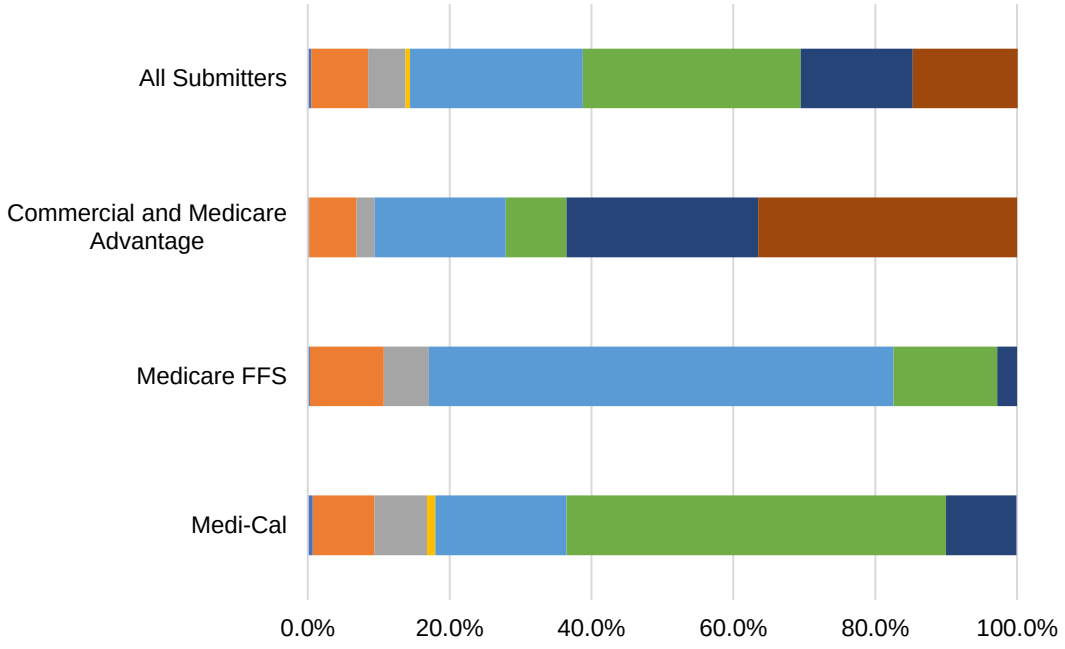
Data from the HPD System for payer-submitted race, ethnicity, and language shows variability by submitter type, as shown in Exhibit 26. These elements are generally well populated in the data supplied by DHCS, for Medi-Cal members. CMS, in their Medicare FFS data, provides data for race but not ethnicity or language. Across all commercial and Medicare Advantage submitters, race, ethnicity, and language were populated 63.5%, 49.4%, and 70.1% of the time.

Exhibit 25. Completeness for Race, Ethnicity, and Language (2021)



More informative than the percent complete is the distribution of values for these elements. Exhibits 26-27 show the distribution of race, ethnicity, and language values in the HPD System for calendar year 2021. The Medicare FFS data is most complete for race, with all but 2.8% of records coded with a value other than Unknown/Not Specified or Missing. The DHCS data is also well distributed and complete for all but 9.9% of the records. By contrast, less than 50% of the commercial and Medicare Advantage data records have a race value. Ethnicity and Language reporting rates are similar, with the DHCS data considerably more complete than the commercial and Medicare Advantage data (Medicare FFS does not include Ethnicity or Language).

Exhibit 26. Distribution of Race Values by Submitter Type (2021)

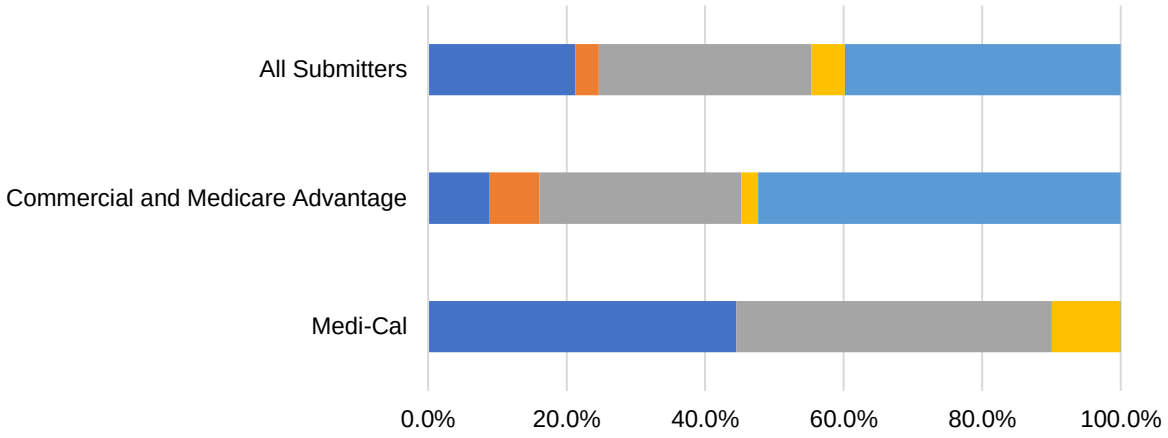


	Medi-Cal	Medicare FFS	Commercial and Medicare Advantage	All Submitters
American Indian/Alaska Native	0.7%	0.3%	0.2%	0.5%
Asian	8.7%	10.4%	6.7%	8.0%
Black/African American	7.4%	6.4%	2.5%	5.3%
Native Hawaiian/Other Pacific Islander	1.2%	0.0%	0.0%	0.6%
White	18.5%	65.5%	18.5%	24.4%
Other	53.5%	14.6%	8.6%	30.7%
Unknown/Not Specified	9.9%	2.8%	27.0%	15.8%
Missing	0.0%	0.0%	36.5%	14.8%

Notes

- DHCS maps data from their systems into a single combined race and ethnicity field into the APCD-CDL™ Race and Ethnicity elements. Records in the DHCS system with Hispanic values are mapped to "Other."
- For more information about race reporting in Medicare, see the OIG Issue Brief *Inaccuracies in Medicare's Race and Ethnicity Data Hinder the Ability To Assess Health Disparities.*
- Percentages based on percentage of eligibility records. Since one individual may have multiple eligibility records for different types of coverage, the percentages may not reflect the percentage of the covered population.

Exhibit 27. Distribution of Ethnicity Values by Submitter Type (2021)

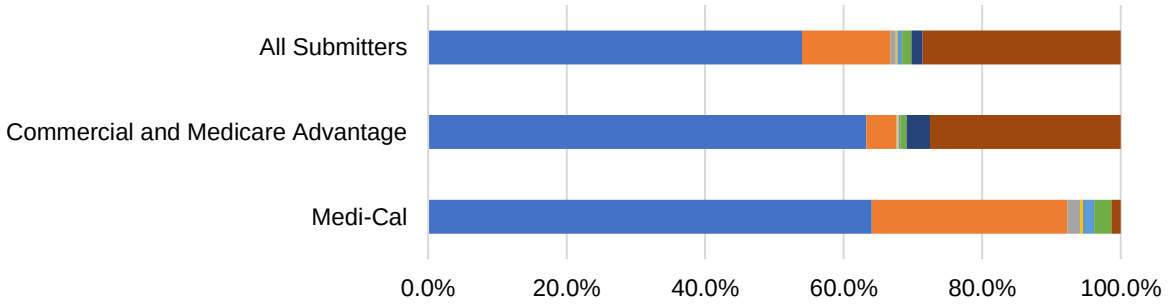


	Medi-Cal	Commercial and Medicare Advantage	All Submitters
Hispanic or Latino	44.5%	8.9%	21.3%
Not Hispanic or Latino	0.0%	7.2%	3.4%
Reported Race Code	45.5%	29.1%	30.7%
Unknown	9.9%	2.4%	4.8%
Invalid or Missing	0.0%	52.3%	39.8%

Notes

- Percentages based on percentage of eligibility records. Since one individual may have multiple eligibility records for different types of coverage, the percentages may not reflect the percent of the covered population.

Exhibit 28. Distribution of Language Values by Submitter Type (2021)



	Medi-Cal	Commercial and Medicare Advantage	All Submitters
English	64.1%	63.3%	54.0%
Spanish	28.3%	4.3%	12.8%
Chinese	1.8%	0.1%	0.7%
Korean	0.4%	0.2%	0.2%
Vietnamese	1.6%	0.2%	0.7%
Other Values	2.5%	0.9%	1.4%
Unknown	0.0%	3.4%	1.6%
Invalid or Missing	1.3%	27.5%	28.6%

Notes

- Percentages based on percentage of eligibility records. Since one individual may have multiple eligibility records for different types of coverage, the percentages may not reflect the percent of the covered population.

There is considerable variability in the completeness of race, ethnicity, and language data among the commercial and Medicare Advantage submitters. Exhibits 29-31 show the number of submitters and percentage of all commercial and Medicare Advantage eligibility records with "actionable" values (i.e., records coded with value other than Unknown, Invalid, or Missing).

- For race, seven of the 36 submitters, representing 27% of the commercial and Medicare Advantage eligibility records, provide relatively complete reporting, with each submitter providing an actionable race value on more than 90% of their records. By contrast, 17 submitters, representing 18% of the records, report an actionable race value less than 10% of the time.
- For ethnicity, four of the 36 submitters, representing 46% of the commercial and Medicare Advantage eligibility records, provide relatively complete reporting, with each submitter providing an actionable value on more than 85% of their records. By contrast, 21 submitters, representing 25% of the records, don't report any actionable ethnicity values.
- For language, 17 of the 36 submitters, representing 46% of the commercial and Medicare Advantage eligibility records, provide relatively complete reporting, with each submitter providing an actionable value on more than 95% of their records. By contrast, nine submitters, representing 24% of the records, don't report any actionable language values.

Exhibit 29. Actionable Race Values Reported by Commercial and Medicare Advantage Submitters, 2021

CATEGORY	NUMBER OF SUBMITTERS	PERCENT OF RECORDS
>90% Actionable Race Values	7	27%
50-89% Actionable Race Values	6	2%
25-49% Actionable Race Values	2	4%
10-24% Actionable Race Values	4	48%
<10% Actionable Race Values	17	18%

Note: "Actionable" values include values other than Unknown/Not Specified, Invalid, and Missing

Exhibit 30. Actionable Ethnicity Values Reported by Commercial and Medicare Advantage Submitters, 2021

CATEGORY	NUMBER OF SUBMITTERS	PERCENT OF RECORDS
>85% Actionable Ethnicity Values	4	46%
50-84% Actionable Ethnicity Values	4	1%
25-49% Actionable Ethnicity Values	3	3%
1-24% Actionable Ethnicity Values	4	25%

0% Actionable Ethnicity Values	21	25%

Note: "Actionable" values include values other than Unknown, Invalid, and Missing

Exhibit 31. Actionable Language Values Reported by Commercial and Medicare Advantage Submitters, 2021

CATEGORY	NUMBER OF SUBMITTERS	PERCENT OF RECORDS
>95% Actionable Language Values	17	46%
90-94% Actionable Language Values	3	23%
40-89% Actionable Language Values	4	3%
1-39% Actionable Language Values	3	4%
0% Actionable Language Values	9	24%

Note: "Actionable" values include values other than Unknown, Invalid, and Missing

The HPD Program actively reviews and discusses data quality with submitters to better understand how the data can be improved over time. HCAI expressed its priority interest in race and ethnicity data at the October 2023 Submitter Group, continued this discussion at the January 2024 Submitter Group meeting, and will conduct individual discussions with submitters starting in the second quarter of 2024. HCAI is in the process of developing individual reports for each submitter to share and review their completion percentages for key data elements, how the completeness rate has trended over time, and how completion compares to other submitters. These reports will also display submission timeliness and variance request trends. The goal of these reports is to provide HCAI a better understanding of limitations submitters may face with providing HPD data elements, to further document how submitters gather, store, and report race, ethnicity, and language data, and to encourage improvement in future reporting. HCAI plans to conduct this review with submitters on an annual basis.

The system-wide categorization and collection of these data by health plans and insurers may need to be addressed at a state-wide or national level. Separate approaches may be required for Medi-Cal, commercial, and CMS-provided Medicare data. The following NCQA recommendations provide a good starting point for consideration for commercial submitters:[29]

- *Specify a set of use cases that explain how race and ethnicity data can be used, and that stipulate permissible and acceptable use from the perspectives of healthcare entities, patients, and community members.*
- *Coordinate a diverse group of stakeholders to develop guidance for implementing interoperability standards that support the collection, use, and sharing of electronic race and ethnicity data for equity reporting.*

Efforts are underway at the state level that will encourage California's health plans to improve the race and ethnicity data in their systems. In response to AB 133 (Committee on Budget, Chapter 143, Statutes of 2021), DMHC established a Health Equity and Quality Measure Set.

Health plans began reporting these measures in 2023, stratified by race and ethnicity.[35] As part of this effort, DMHC will develop a process to track what demographic data health plans have collected and for what percentage of their enrollees. HCAI has a seat on the DMHC Health Equity and Quality Committee, and will continue to work with DMHC to better understand guidance that DMHC is providing to plans regarding data quality/completion and to discuss DMHC interest in using HPD demographic data.

HCAI is an active member of NAHDO, and HCAI staff sit on the NAHDO Board of Directors. Beginning when HCAI first obtained authority to implement the HPD Program in 2018, and through consistent engagement with state and national partners, HCAI has been committed to promoting the adoption of national standards that meet the needs of California, including the uniqueness of California's diverse communities. HCAI has been a staunch and early advocate of expanding the APCD-CDL™ categories for race and ethnicity, and sexual orientation and gender identity. Through collaborative effort among states on the APCD Council, Version 3 of the APCD-CDL™ for the first time enables APCDs to collect granular breakouts for race and ethnicity, and provides health plans the ability to record new categories for the sexual orientation and gender identities of their enrollees. The NAHDO APCD Council voted and ratified these changes in September 2022.

HCAI will continue to advocate for changes to standards and processes that improve the ability to analyze disparities in access and health outcomes by collaboratively participating with state and national partners in alignment with California's needs and values.

While efforts continue to improve collection of race and ethnicity data according to current standards, the federal government is considering changes to the standards. In January 2023, the Office of Management and Budget (OMB) published initial proposals for updating race and ethnicity statistical standards. Changes under consideration include: collect race and ethnicity information using one combined question, add "Middle Eastern or North African" as a new minimum category, and require the collection of detailed race and ethnicity categories by default.[36]

HCAI will continue to discuss public reporting priorities with the HPD Advisory Committee, including initial reporting on race, ethnicity, and language, while abiding with the HPD Program's reporting principles, including ensuring that HPD data analysis is feasible and credible given the quality of the data. In October 2023, the HPD Advisory Committee included social drivers of health as a public reporting priority for HCAI in 2024.

Encounter Data Completeness in the HPD

Capitation and other alternative payment arrangements are common in California's market. As described above (see also inset, How do Claims differ from Encounters in Health Care?) administrative data includes both paid claims and no-pay encounters. The inclusion of encounter data in the HPD is essential to provide a more complete picture of health care services received by patients. Yet, given the lack of direct incentive to submit these services for payment, there has long been a concern that encounters are under-reported. Complex relationships between provider organizations and upstream systems compound the data reporting and collection challenge, with many data handoffs and potential loss of data along the way.

Like other aspects of data quality, the HPD Program benefits from efforts by others to improve encounter data completeness, including efforts by DHCS, CMS, Covered California, and California's Encounter Data Improvement Program.

DHCS has long required health plans to submit complete and timely encounter data, and sponsors annual encounter data validation studies that include a review of medical records and measures the completeness and accuracy of the professional encounter data submitted to DHCS.[37]

CMS collects encounters from Medicare Advantage plans and uses a measurement framework that includes an annual evaluation of encounter data performance metrics for Medicare Advantage contracts. The threshold for each metric is designed to identify performance that is below reasonable expectations, and plans falling below the threshold are subject to compliance action.[38] Even though the HPD Program collects Medicare Advantage data directly from California's plans and insurers (rather than from CMS), the quality of the data likely benefits from this CMS focus on encounter data completeness.

How do Claims differ from Encounters in Health Care?

CLAIM	ENCOUNTER
Records that a service was provided.	
Shows service, diagnosis, patient, and provider information.	
Requests payment for services. Claims processor may accept, reject, or return for correction.	Does not request payment.
Incentive to submit is payment. Other incentives may include risk-sharing, quality performance, and/or other financial incentives	Weaker incentive to submit because does not generate payment. Incentives to submit may include compliance requirements, quality incentives, and serving as the basis for future years' compensation.

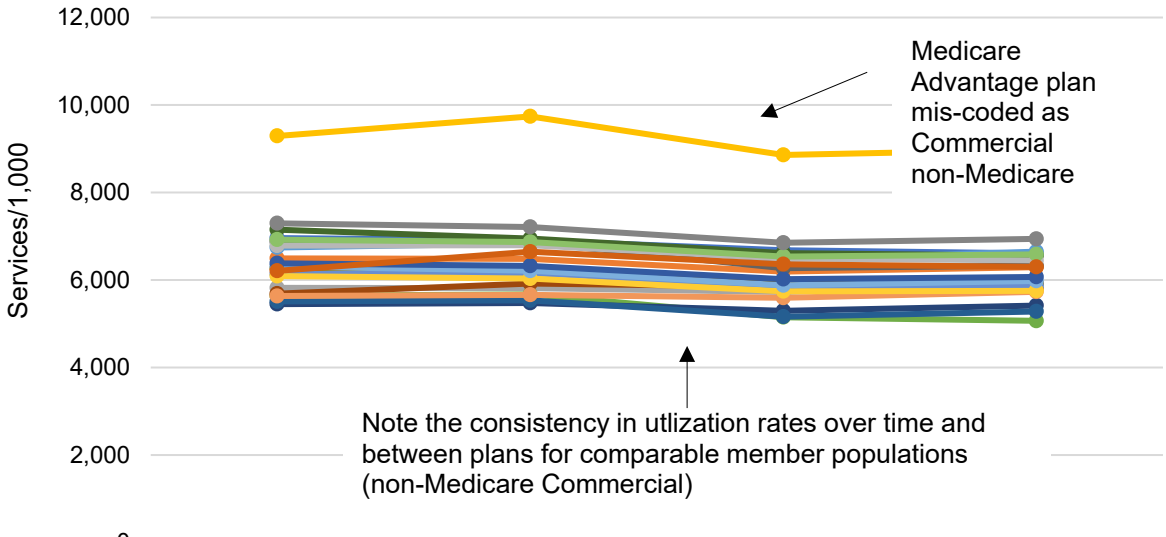
One of the requirements of the 2016 merger between Centene and Health Net imposed by the DMHC was investment of $50 million, some of which was targeted to improve the completeness and accuracy of encounter data.[39] The Encounter Data Improvement Program (EDIP) has launched several activities, including an encounter data market research study, one-time improvement funding, provider-level assessment and implementation grants, a stakeholder engagement process, and established governance.[40] In 2021, $26 million of EDIP's budget was committed to establish oversight of encounter data improvement efforts in California, and IHA

was selected to oversee, monitor and implement encounter data improvement efforts, including:[40]

- Establish a statewide governance body to develop and organize policies, standards, and processes
- Develop standardized systems and coding for data submissions
- Develop encounter data training sessions, technology support, and technical assistance programs for providers and staff

Exhibit 32 displays services per 1,000 commercial (non-Medicare) members by year and health plan. This type of analysis, comparing data submitted by multiple submitters over time, provides an indication—although not definitive—of encounter data completeness. The exhibit also exemplifies the importance of this type of analysis for evaluating data quality. The outlier (labeled Plan 4) is a Medicare Advantage plan mis-labeled as a commercial plan (the submitter has been notified and has subsequently corrected the data). Medicare members use more services than commercial non-Medicare members, so the utilization rate for Plan 4 is understandably higher and similar to the rates for Medicare populations. The results for the non-Medicare plans are encouraging from an encounter data completeness perspective, given the consistency in utilization rates over time and across plans, despite differences among types of plans in terms of the use of capitated services. A more robust assessment of encounter data completeness would require audits of provider systems and patient charts against submitted encounter data, which is beyond the scope of this report.

Exhibit 32. Services per 1,000 Commercial Non-Medicare, by Health Plan, 2020



	2018	2019	2020	2021
Plan 1	6,959	6,907	6,679	6,609
Plan 2	6,494	6,482	6,192	6,292
Plan 3	5,821	5,814	5,727	5,923
Plan 4	9,292	9,739	8,858	8,978
Plan 5	6,743	6,836	6,483	6,648
Plan 6	5,637	5,671	5,149	5,067
Plan 7	5,449	5,474	5,296	5,415
Plan 8	5,689	5,919	5,848	5,791
Plan 9	6,875	6,859	6,270	6,328
Plan 10	6,266	6,166	5,938	5,916
Plan 11	5,509	5,535	5,161	5,281
Plan 12	7,151	6,947	6,616	6,556
Plan 13	6,132	6,163	5,854	5,862
Plan 14	5,634	5,661	5,594	5,730
Plan 15	6,785	6,795	6,490	6,448
Plan 16	6,090	6,033	5,741	5,752
Plan 17	6,311	6,211	5,879	5,980
Plan 18	6,922	6,868	6,535	6,586
Plan 19	6,382	6,323	6,019	6,071
Plan 20	6,210	6,647	6,359	6,300
Plan 21	7,295	7,212	6,851	6,940

Other Areas of Data Quality

For the purposes of this report, not all areas of data quality were evaluated. For example, payment fields such as paid amount, allowed amount, deductible, coinsurance, co-payment, and premium will be important for many analyses using the HPD but are still under evaluation by HCAI and are not included in this report. HCAI will continue to evaluate data quality and share the results with data users and other stakeholders.

Known Data Quality Issues

The HPD Program employs robust data quality evaluation and improvement processes and is committed to improving the quality and usefulness of the information in the HPD System over time. The HPD Program will also continue to document and communicate information about the data in the HPD System to end-users (for examples, see the technical notes for the HPD Snapshot and the HPD Healthcare Measures). The items below represent an initial list of known data quality issues specific to the HPD System.

1. ***Data Collection—Years of Data Not Included.*** Depending on when users access the data in the HPD System, data for certain payers may not be available. For example, Medicare FFS data for 2021 were not available and were absent from the initial HPD Snapshot released in June 2023. Documentation for the HPD Snapshot and other public reports will continue to explain any such exclusions.
2. ***Count of unique individuals***. The HPD System, like all APCDs, uses a master person index process to match members and patients across different payers and over time. The process uses data collected from payers such as name, date of birth, sex, health plan IDs, and social security numbers, when available. The HPD Program intends to continuously refine and improve its approach to matching; such improvements may increase the cases in which two or more records are determined to belong to a single individual and therefore reduce the count of unique individuals represented in the HPD System.
3. ***Identification of members enrolled in ERISA self-funded plans***. Measuring the number of self-funded lives in the HPD is challenged by data issues, including variability in how submitters use the available data elements in their submissions. Although the APCD-CDL™ includes an element for submitters to report self-funded status, there is no straight-forward way to differentiate ERISA vs. non-ERISA self-funded plans and lives.
4. ***County address used as home address for some Medi-Cal members***. In some cases, the home address for Medi-Cal members is listed as the address of the county building or other agency, resulting in many members with the same address in the HPD System. This most commonly occurs for homeless beneficiaries and for those who state they do not have a fixed address; counties use established standards for completion of the residence address, including the address of the county building or other agency where mail is held for the beneficiary.[41]
5. ***Diagnosis data missing on some data from DHCS***. Approximately 6% of DHCS records submitted for medical services are missing a principal diagnosis code. More than half of these records come to DHCS from the Department of Developmental Services (for the Case Management program) and from the Department of Social

Services (DSS) (for the Personal Care Services Program). None of the records for those two programs have diagnosis codes.

6. ***Provider identifiers and analysis of provider information***. Due to variability in the use of provider identifiers and other data elements used to determine the facility, provider group practice, and setting of care, analysis attributing information to providers is complex. For example, the number of services provided at specific facilities in the HPD System does not always compare well to external sources. It is expected improvements to provider data sources and HCAI's ability to match information about providers and settings of care will improve over time.

6. Public Reporting and Data Release

Public Reporting

This section provides an update on activities related to making data from the HPD Program available to the public, researchers, and others. The information from the HPD System is intended to support greater health care cost transparency to inform policy decisions regarding the provision of quality health care, and to reduce health care costs and disparities. It is also intended for the information to be used to develop innovative approaches, services, and programs that may have the potential to deliver health care that is both cost effective and responsive to the needs of all Californians.

HCAI has released several public data products, including downloadable files with thousands of de-identified rows of data. The Healthcare Payments Data (HPD) Snapshot[3] provides an overview of data currently available as submitted in the HPD System with visualizations that allow users to explore how many Californians received coverage from each type of payer and the number of medical or pharmacy service records generated. The Healthcare Payments Data (HPD) Measures[4] allow users to explore the care and characteristics of Californians within the HPD System across three measurement categories: health conditions, utilization, and demographics. The Snapshot and Measures visualizations allow users to apply filters and grouping options, and users can download the detailed data, subject to the California Health and Human Services Agency's Data De-Identification Guidelines.[8]

These products are being used by various stakeholders to inform healthcare decisions, and HCAI continues to evaluate and monitor the use of HCAI data. Additional data products will be developed and released over time, with input from and considering the requests of stakeholders, that address the most pressing and important issues of health policy. Public reporting priorities for 2024 discussed with the HPD Advisory Committee at the October 2023 meeting include:

1. ***Health equity***: evaluate differences in services and outcomes between population groups and associate healthcare measures and costs with social drivers of health.
2. ***Hospital sector spending***: report utilization and costs for inpatient hospital services and display hospital-based episodes of care.
3. ***Enhanced pharmaceutical sector spending***: in addition to reporting the top 25 drugs from a cost, volume, and patient out-of-pocket costs perspective, consider displays on the relationship between wholesale and final cost, variations by region or payer type, and evaluations of specific drug types, such as biologics or opioids.

HCAI will continue to produce new public analytic reports and update existing reports regularly.

HCAI anticipates continuing to advance the accessibility and usefulness of HPD public reports as the database becomes more comprehensive and complete and HCAI builds its capacity over time. As the quality and completeness of HPD data quality improves, so will the potential for producing more analyses with the HPD. The HPD Advisory Committee continues to advise HCAI on public reporting priorities, and has acknowledged that more complex, and potentially

informative, uses of HPD data will grow over time, such as studying episodes of care, comparisons of cost and quality, and provider networks and payment arrangements.

Data Release

In addition to continuing to expand upon public reporting, HCAI is in the process of establishing a data release program for non-public data and intends to begin considering requests for data in 2024. Different from HPD data available to the public, which is aggregated and de-identified, non-public data may include personally identifiable information or other sensitive information about patients or individual consumers. The data release program is critical to providing researchers, and other eligible requestors of HPD data, controlled access to such information to support longitudinal and population level analyses of California's healthcare system. A core tenet of the HPD Program is protecting individual privacy and safeguarding access to sensitive data within the HPD System. (Read more about how HCAI data is protected.[42])

The HPD statute requires HCAI to establish, through regulations, a "data use, access, and release program" to provide HPD data to outside entities while protecting privacy (HSC, § 127673.82 subd. [b]). The statute states how data can be made available to members of the public and other state agencies and provides "privacy protection standards" that can be supplemented through these regulations (HSC, § 127673.83). For non-public data, the HPD statute also requires approvals from one or two state committees, the HPD Data Release Committee and the California Committee for the Protection of Human Subjects (HSC, § 127673.83 subd. [b][2][c]). DHCS must review and approve releases of non-public data for Medi-Cal and other DHCS-administered programs. The law also requires HCAI to prioritize use of a secure research environment for access to non-public data (HSC, § 127673.82).

HCAI established the HPD Data Release Committee (DRC) in December 2022 to advise the HPD Program on policies and procedures for access to non-public HPD data. The DRC is made up of subject matter experts representing key stakeholder groups including consumer advocates, labor, providers (both at clinician and hospital levels), payers, purchasers, suppliers, and researchers. Members have deep knowledge and experience with health care data, privacy, and security (see Exhibit 33.E for more information on the DRC). The DRC also has the authority to provide advice to the HCAI Director on privacy and security matters related to the HPD Program and provide feedback to the department on the data application and review process.

Activities of the DRC to date have focused on creating the policies and procedures of the data release program, which must meet goals related to public benefit of broad use of the data as well as protection of patient privacy. Once the program begins accepting applications for non-public HPD data, the DRC will review and make recommendations to HCAI on access to and release of these data, considering whether the use of the data is consistent with the goals of the HPD Program including: (1) whether it provides greater transparency regarding health care costs, utilization, quality, or equity, or (2) how the information may be used to inform policy decisions regarding the provision of quality health care, improving public health, reducing health disparities, advancing health coverage, or reducing health care costs.

Regulations to create the HPD's Data Use, Access, and Release Program, and implement the statutory requirements discussed above were submitted to the Office of Administrative Law and with the initial public comment period in the summer of 2023. Based on comments received during the initial public comment period, HCAI proposed changes and a second round of public comments ran from December 18, 2023, to February 1, 2024. HCAI intends to finalize the regulations in the first quarter of 2024. In the interim, HCAI continues to work with the DRC on procedures for evaluating and managing requests and developing and testing the data request and data access systems. All DRC meetings are open to the public, and public testimony from data users, other stakeholders, and the public, is welcome.

HCAI also continues to actively explore avenues to provide access to Medicare FFS data that are less burdensome for researchers. The HPD Program obtains Medicare FFS and Prescription Drug Program data from CMS through a state agency request process, similar to most other state APCDs. The Data Use Agreement (DUA) between CMS and states allows APCDs to integrate the CMS data with other data, conduct analyses, produce aggregate reports, and share the CMS data with other state agencies conducting research. It does not, however, allow HCAI to disseminate the Medicare FFS data to external users of the HPD, such as academic researchers. HCAI is working with CMS, and other partners, to consider ways to provide appropriate access to such data on the Medicare population to the HPD research community.

7. Findings and Next Steps

In June 2023, The Department of Health Care Access and Information (HCAI) released the Health Care Payments Data (HPD) Program's first public data, marking the completion of initial design, development, and implementation of California's All-Payer Claims Database (APCD). The HPD System, already the nation's largest APCD, provides an unprecedented opportunity to understand and address healthcare costs and to drive improvement in California's healthcare system. Based on progress to date, HCAI is well-positioned to fully realize the intent outlined by the Legislature (HSC, § 127671-127674):

- Establish a system to collect information regarding the cost of health care and a process for aggregating such information from many disparate systems, with the goal of providing greater transparency regarding health care costs.
- Improve data transparency to achieve a sustainable health care system with more equitable access to affordable and high-quality health care for all.
- Encourage use of such data to deliver health care that is cost effective and responsive to the needs of enrollees, including recognizing the diversity of California and the impact of social determinants of health.

A summary of key findings and next steps is presented below.

1. ***California's APCD was completed on time***. HCAI released the HPD Program's first public data in June 2023, including summary enrollment and healthcare utilization information for more than 30 million Californians, for calendar years 2018 through 2021. Publication of the Healthcare Payments Data (HPD) Snapshot[3] marked the successful culmination of a multi-year effort of legislation, planning, data collection, and implementation of California's APCD. Release of the Snapshot data also satisfied the legislative requirement that the development of the HPD System "be substantially completed" no later than July 1, 2023 (HSC, § 12671). The HPD will continue to add data years to analytic extracts and reports and is collecting data monthly.

2. ***The HPD System includes all the initially planned data types, sources, and time periods*.** The HPD System includes data from all the planned sources in the state (see the 2020 Health Care Payments Data Program Report to the Legislature[1]), including all Medi-Cal and Medicare FFS covered lives and all covered lives from California's health plans and insurers subject to the reporting mandate from 2018 forward, including, for calendar year 2021:
    - 16.8 million non-Medi-Cal members from California's health plans and insurers, including commercial and Medicare Advantage
    - 14.1 million Medi-Cal members, including 11.7 million in managed care plans and 2.4 million in Medi-Cal FFS
    - 3.4 million members in Medicare FFS

3. ***The HPD System reflects approximately 82% of Californians and their healthcare services.*** The HPD System includes services and eligibility records for approximately 31.5

million individuals with medical coverage in one or more healthcare plans for each reporting year. Using comparison data, the HPD System includes approximately:

- Member information for 82% of California's total population and 89% of California's insured population
- 90% of state-wide ED visits
- 85% of inpatient admissions
- 76-89% of office visits

In addition, nearly all of California's providers are represented in the HPD System, including over 83,000 individual physicians.

4. ***Efforts to expand the HPD are already underway, including adding data from dental plans and insurers, capitation payments and other non-claims payment data***. Dental data collection will begin in 2024 and non-claims payment data in 2025.

5. ***Increasing voluntary data from private self-funded arrangements provides the biggest opportunity to increase the content and generalizability of the HPD.*** Preliminary analysis of the self-funded lives in the HPD indicates that voluntary participation of ERISA plans is low and that as many as 3.2 million ERISA self-funded lives are not yet included in the HPD. HCAI plans to conduct additional targeted outreach to large employers and other purchasers to encourage voluntary submission. State policymakers should consider policy changes that encourage participation, such as requiring ASOs to provide an opt-in form to their clients or policies that limit the fees ASOs are able to charge to submit data to the HPD.

6. ***Collecting data directly from providers and suppliers on a limited basis could prove useful but would add considerable cost and complexity to the operation of the HPD Program.*** Adding provider and supplier organizations to the list of required submitters, or allowing them to submit voluntarily, assuming they were to submit the same types of data files currently required of plans and payers, would exponentially increase the number of submitters and files and require new efforts to find and eliminate duplicate services and payments records. Limiting the effort to payments made by Risk Bearing Organizations to their downstream contracting partners could prove more useful, but significant technical challenges, feasibility questions, and resource considerations would need to be addressed.

7. ***Preliminary analyses of data quality indicate that the data quality in the HPD System is reflective of and consistent with administrative data used in healthcare operations, and there are opportunities for improvement, particularly for demographic data.*** By its nature, administrative data is not originally intended for use by researchers in analytic databases such as an APCD, but it has proven to provide rich analytic value and represents the most accessible source for the detailed healthcare services and payments provided in a healthcare system. While required data fields are complete and accurate, collection of some demographic data is lacking and can be improved.

HCAI is committed to continually improving the quality and value of the data within the HPD System. HCAI will share data quality results with submitters and discuss strategies for improvement. HCAI may update benchmarks and revise criteria for selected data elements, based on existing data submissions as well as data from other states, to raise the bar for data quality and completeness. HCAI is also working with DMHC and other agencies on coordinated efforts to improve the collection, storage, and submission of standardized race and ethnicity and other critical elements of data. Additionally, the use of data from the HPD System will further inform efforts to improve data quality and make the quality of data transparent to data users and other stakeholders.

8. ***HCAI's strong partnership with NAHDO and influence on national standards has greatly benefited the HPD Program.*** As a health data organization, HCAI has long-standing leadership and expertise in administrative healthcare data and associated standards and specifications. This, and HCAI's recognized influence on the national level, has been an instrumental part of HCAI's success on the HPD Program.

9. ***The HPD Program's public reporting and data release functions have been successful and continue to evolve.*** The public information HCAI has produced from the HPD Program already represents a significant expansion in the availability of actionable, transparent healthcare data in California, and continued development of a data use, access, and release program will provide avenues for researchers and others to securely access non-public HPD data.

10. ***HCAI has previously made recommendations to state policymakers to fully fund the HPD Program long-term.*** In March 2023, HCAI submitted a report to the legislature on the long-term funding options for the HPD Program[2]. In that report, HCAI made the following recommendations:
    - Support an annual total funds budget of $22 million for the HPD Program.
    - Establish a state funding model, using General Fund, special funds, or some combination thereof, that provides $15.4 million in annual state funds.
    - Ensure the above funding provisions are in place with Fiscal Year 2025-26 to avoid disruption to HPD Program operations.

HCAI looks forward to working with all stakeholders to continue to make progress in advancing transparency in health care through data and fulfilling the statutory intent and goals of the HPD Program.

Additional Resources about HCAI and the HPD Program

The links below provide additional information about the California Department of Health Care Access and Information and the Health Care Payments Data (HPD) Program:

Exhibit 33. Additional HCAI and HPD Program Resources

RESOURCE	DESCRIPTION
A. California Department of Health Care Access and Information (HCAI)	General information about HCAI and its programs.
B. Health Care Payments Data Program	Information on the HPD Program, including news, goals, FAQs, stakeholder engagement, published data, and upcoming activities.
C. HPD Program Advisory Committee	Information about the HPD Program Advisory Committee, including purpose, membership, and past and future meeting materials.
D. HPD Program Data Submitters	Information for HPD Submitters, including the Data Submission Guide and other resources, FAQs, and past and future meeting materials.
E. HPD Program Data Release Committee	Information about the HPD Program Data Release Committee, including purpose, membership, and past and future meeting materials.
F. Healthcare Payments Data Program: Voluntary Submitters	Information about voluntary submission to the HPD Program, including potential benefits to employers, FAWs, and an opt-in form.
G. 2020 Legislative Report: Health Care Payments Data Program Report to the Legislature	Includes background and learnings from other state APCDs, as well as 36 specific recommendations, discussed and voted on by Review Committee members, for the successful operation of the HPD Program in California, across nine areas: • APCDs and Use Cases • Data Categories and Formats • Linkages • Submitters • Funding and Sustainability • Privacy and Security • Technology Alternatives • Data Quality • Governance
H. 2023 Legislative Report: Long-Term Funding Options For The Health Care Payments Data Program	Summarizes long-term funding options for the program, for consideration by the legislature

References

1. California Department of Health Care Access and Information. *Health Care Payments Data Program, Report to the Legislature*.; 2020. Accessed December 17, 2023. https://hcai.ca.gov/wp-content/uploads/2020/12/HPD-Legislative-Report-20200306.pdf

2. California Department of Health Care Access and Information. *Long-Term Funding Options for the Health Care Payments Data Program*.; 2023. Accessed September 1, 2023. https://hcai.ca.gov/wp-content/uploads/2023/02/Funding-Report-March-2023-1.pdf

3. California Department of Health Care Access and Information. Healthcare Payments Data (HPD) Snapshot. Accessed September 1, 2023. https://hcai.ca.gov/visualizations/healthcare-payments-data-hpd-snapshot/

4. California Department of Health Care Access and Information. Healthcare Payments Data (HPD) Measures: Health Conditions, Utilization, and Demographics. Accessed September 12, 2023. https://hcai.ca.gov/visualizations/healthcare-payments-data-hpd-measures-health-conditions-utilization-and-demographics/

5. APCD Council. APCD Legislation by State. Published 2022. Accessed September 1, 2023. https://www.apcdcouncil.org/apcd-legislation-state

6. U.S. Food and Drug Administration. FDA Drug Label for Ventolin Hfa. Accessed October 25, 2023. https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/020983s032lbl.pdf

7. U.S. Food and Drug Administration. FDA Drug Label for ELIQUIS. Accessed October 25, 2023. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/202155s000lbl.pdf

8. California Health and Human Services. *Data De-Identification Guidelines*.; 2016. Accessed December 17, 2023. https://chhsdata.github.io/dataplaybook/documents/CHHS-DDG-V1.0-092316.pdf

9. California Health and Human Services Open Data Portal. Medi-Cal Certified Eligibles Tables, by County from 2010 to Most Recent Reportable Month. Accessed October 25, 2023. https://data.chhs.ca.gov/dataset/medi-cal-certified-eligibles-tables-by-county-from-2010-to-most-recent-reportable-month/resource/2c28bf78-a385-4d0c-88d5-7d1eef09a5ab

10. Total Number of Medicare Beneficiaries by Type of Coverage. Kaiser Family Foundation. Accessed December 9, 2023. https://www.kff.org/medicare/state-indicator/total-medicare-beneficiaries/

11. Wilson K. *California Health Insurers, Enrollment — 2023 Edition*.; 2023. Accessed December 7, 2023. https://www.chcf.org/publication/ca-health-insurers-enrollment-2023-edition/

12. What Is VEBA? California Schools VEBA. Accessed January 9, 2024. https://www.vebaonline.com/about/

13. Macrotrends. California Population 1900-2022. Accessed September 19, 2023. https://www.macrotrends.net/states/california/population

14. U.S. Census Bureau. QuickFacts: California. Accessed September 19, 2023. https://www.census.gov/quickfacts/fact/table/CA/PST045222

15. National Plan & Provider Enumeration System. Taxonomy Page -- NPPES. Accessed January 10, 2024. https://nppes.cms.hhs.gov/webhelp/nppeshelp/TAXONOMY%20PAGE.html

16. National Provider Identifier (NPI) May 23, 2008, Implementation. Accessed January 10, 2024. https://www.cms.gov/newsroom/fact-sheets/national-provider-identifier-npi-may-23-2008-implementation

17. APCD Council. SCOTUS Gobeille v. Liberty Mutual Insurance Company Decision. APCD Council. Accessed January 2, 2024. https://www.apcdcouncil.org/resources/public-comments/scotus-gobeille-v-liberty-mutual-insurance-company-decision

18. National Association of Insurance Commissioners. Health and Welfare Plans Under the Employee Retirement Income Security Act: Guidelines for State and Federal Regulation. Published online 2018. Accessed August 30, 2019. https://content.naic.org/sites/default/files/inline-files/cmte_b_erisa_exposure_erisa_handbook_chair_draft_revisions.pdf

19. U.S. Department of Labor. *Health Insurance Coverage Bulletin 2021*.; 2022. Accessed September 7, 2023. https://www.dol.gov/sites/dolgov/files/EBSA/researchers/data/health-and-welfare/health-insurance-coverage-bulletin-2021.pdf

20. Utah Office of Administrative Rules. Voluntary Opt-In for a Self-Funded Plan (R590-262-7). Accessed January 9, 2024. https://adminrules.utah.gov/public/rule/R590-262/Current%20Rules

21. California Department of Managed Health Care. SB 260 Capitated Providers June 2023. Accessed October 31, 2023. https://www.dmhc.ca.gov/Portals/0/Docs/OFR/RBOs/SB%20260%20Capitated%20Providers%20June%202023.pdf?ver=SN3SoGLRhzDb0pPtLqAmYw%3d%3d

22. California Health and Human Services Open Data Portal. Hospital Annual Utilization Report & Pivot Tables. Accessed October 31, 2023. https://data.chhs.ca.gov/dataset/hospital-annual-utilization-report

23. Definitive Healthcare. Top 25 California physician groups by Medicare charges. Accessed October 30, 2023. https://www.definitivehc.com/resources/healthcare-insights/top-california-physician-groups-medicare-charges

24. California Health Care Foundation. *California Physicians, 2021: A Portrait of Practice*.; 2021. https://www.chcf.org/wp-content/uploads/2021/03/PhysiciansAlmanac2021.pdf

25. California Department of Health Care Access and Information. Economic and Fiscal Impact Statement, Health Care Payments Data Program. Published online July 31, 2023.

26. California Department of Health Care Access and Information. HPD Data Submission Guide 2.0. Published July 17, 2023. Accessed December 21, 2023. https://hcai.ca.gov/wp-content/uploads/2023/12/HPD-Data-Submission-Guide-2.0_Final.pdf

27. Department of Health Care Services. Medi-Cal Managed Care Encounter Data Reporting. Accessed September 26, 2023. https://www.dhcs.ca.gov/dataandstats/data/Pages/MMCDClmsEncDataRpt.aspx

28. Medical Coding Modifiers - CPT®, NCCI & HCPCS Level II - AAPC. Accessed September 30, 2023. https://www.aapc.com/resources/what-are-medical-coding-modifiers

29. Soucie J, Buckley RO, Albanese K, Harrington R, Scholle SH. *Current Health Plan Approaches to Race and Ethnicity Data*.; 2023. https://www.ncqa.org/wp-content/uploads/2023/03/Current-Health-Plan-Approaches-to-Race-and-Ethnicity-Data-Collection-and-Recommendations-for-Future-Improvements_Final.pdf

30. Soucie J, Scholle SH, Ayo-Vaughan M, et al. How Health Plans Could Improve the Collection of Race and Ethnicity Data. Published April 3, 2023. Accessed January 2, 2024. https://www.commonwealthfund.org/blog/2023/how-health-plans-could-improve-collection-race-ethnicity-data

31. Medicaid and CHIP Payment and Access Commission. *Report to Congress on Medicaid and CHIP, Chapter 1: Medicaid Race and Ethnicity Data Collection and Reporting: Recommendations for Improvement*.; 2023. https://www.macpac.gov/wp-content/uploads/2023/03/Chapter-1-Medicaid-Race-and-Ethnicity-Data-Collection-and-Reporting-Recommendations.pdf

32. U.S. Department of Health and Human Services Office of Inspector General. *Inaccuracies in Medicare's Race and Ethnicity Data Hinder the Ability To Assess Health Disparities*.; 2022. https://oig.hhs.gov/oei/reports/OEI-02-21-00100.pdf

33. National Association of Health Data Organizations. *Current and Innovative Practices in Data Quality Assurance and Improvement*.; 2019. Accessed September 11, 2023. https://www.nahdo.org/sites/default/files/2019-10/Current%20and%20Innovative%20Practices%20in%20Data%20Quality%20Assurance%20and%20Improvement%20.pdf

34. The Center for Improving Value in Health Care (CIVHC). Plaintalk Blog: Collecting Race and Ethnicity Data in the CO APCD. Published December 4, 2021. Accessed October 3, 2023. https://civhc.org/2021/12/03/plaintalk-blog-collecting-race-and-ethnicity-data-in-the-co-apcd/

35. California Department of Managed Health Care. APL 22-028 - Health Equity and Quality Measure Set and Reporting Process. Published online December 21, 2022. Accessed October 5, 2023. https://www.dmhc.ca.gov/Portals/0/Docs/OPL/APL%2022-028%20-%20Health%20Equity%20and%20Quality%20Measure%20Set%20and%20Reporting%20Process%20(12_21_2022).pdf

36. Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards. Federal Register. Published January 27, 2023. Accessed November 28, 2023. https://www.federalregister.gov/documents/2023/01/27/2023-01635/initial-proposals-for-updating-ombs-race-and-ethnicity-statistical-standards

37. California Department of Health Care Services. *2021 – 22 Encounter Data Validation Study Report*.; 2023. https://www.dhcs.ca.gov/dataandstats/reports/Documents/MMCD_Qual_Rpts/2021-22-EncounterDataValidation-Report.pdf

38. The Medicare Payment Advisory Commission. *Report to the Congress: Medicare and the Health Care Delivery System*.; 2019. Accessed September 26, 2023. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/jun19_ch7_medpac_reporttocongress_sec.pdf

39. Centene Corporation. Undertakings Related to Health Net and Centene Merger. Published online March 18, 2026. Accessed October 1, 2023. https://www.dmhc.ca.gov/Portals/0/AbouttheDMHC/NewsRoom/u032216.pdf

40. HealthNet. Encounter Data Improvement Program: Improving Medi-Cal Care & Reducing Health Disparities. Accessed October 1, 2023. https://iha.org/wp-content/uploads/2021/03/HN-Encounter-Data-Program-One-Sheet-FINAL-2021.pdf

41. Department of Health Care Services. Medi-Cal Eligibility Division Information Letter No.: I 16-06. Published online June 30, 2016. https://www.dhcs.ca.gov/services/medi-cal/eligibility/letters/Documents/I16-06.pdf

42. California Department of Health Care Access and Information. How is HCAI data protected? Accessed January 11, 2024. https://hcai.ca.gov/data/data-resources/