

Healthcare Payments Data Program (HPD) Webinar January 2026 – Q&A Summary

Below is a summary of the questions and answers discussed during HCAI's webinars held in January 2026. Also included are questions received live during the webinar, and many that HCAI was unable to answer in the time allotted. HCAI looks forward to seeing you at the next event! [Subscribe to HCAI's mailing list](#) to be notified of events and public meetings.

The Healthcare Payments Data Program (HPD) is California's All-Payer Claims Database or APCD. The HPD is a research database comprised of healthcare administrative data, claims, and encounters, generated by transactions among payers and providers on behalf of insured individuals. The webinars included an overview of the HPD program and updates on data collection, data and reporting, HPD program releases, product demonstrations, and data access and usage. Visit the [HPD webpage](#) for more information.

Any additional questions or feedback may be sent to DataAndReports@hcai.ca.gov.

Data Products

- 1. Are the briefs intended to be point-in-time, or will they continue to be updated as new years of data come in?**

HCAI uses stakeholder and Advisory Committee input, business intelligence, and our annual strategic planning process to identify how HCAI resources will be used to update or retire existing data products. Based on this input, business intelligence, and strategic planning, HCAI will update Data Briefs that are identified as beneficial to refresh.

- 2. Are there upcoming analyses of social drivers of health that the HPD team plans to do?**

HCAI is currently working on analyzing and enriching data fields that will be used in social drivers of health (SDoH) analyses. HCAI is also working to obtain a license for the Healthy Places Index (HPI), which will be used in our SDoH analyses. HCAI is in the early stages of analyzing the preferred languages spoken.

Data Collection

3. What is the exact percentage of records where race and ethnicity are listed as “missing” or “unknown”?

Please refer to the [HPD Data Completeness Fact Sheet](#) pages 3 and 4. The completeness rates for race and ethnicity are published for Commercial, Medicaid, Medicare Fee-For-Service (FFS), and Medicare Advantage. Completeness rates for race, ethnicity, and language are based on actionable fields, which exclude any values of “unknown”, “invalid”, and “missing”.

4. Do social determinants data include sexual orientation and gender identity, and if not, will this be something requested of plans in the future?

HCAI's work on SDoH is in early-stage development. HCAI is currently working to address limited data completeness for key measures, such as race/ethnicity, by developing methods to enrich these data. HCAI is also in the process of obtaining the license to use the HPI. With HPI, HCAI will be looking at SDoH at various geographic levels. Additionally, HCAI is working on an analysis of preferred spoken language.

5. What challenges and/or opportunities are there for enhancing transparency and representation of diverse patient populations and communities with the ongoing effort to update California's standards for demographic data disaggregation, and in accordance with recent changes to the US Office of Management and Budget's (OMB's) Statistical Policy Directive No.15 for Census categories and reporting?

All-Payer Claims Databases, like the HPD, collect administrative data that was originally intended for billing purposes. The HPD gathers race and ethnicity data from plans and insurers, but it is often incomplete, particularly for the commercial payers. The HPD uses the APCD-Common Data Layout (CDL), which is a data layout specification created by the National Association of Health Data Organizations (NAHDO). The APCD-CDL follows the CDC race and ethnicity code set for capturing race and ethnicity as separate data elements. With updates to the OMB Statistical Policy Directive 15, the CDC code set will be revised to capture a single data element for race and ethnicity. We anticipate the NAHDO will, in turn, update the APCD-CDL. In the future, HPD data will align with disaggregation, though data completeness issues may persist.

6. At what level of granularity does the HPD collect and report demographic data for patients?

The member eligibility file includes demographic and plan data for each member with medical, pharmacy, or dental coverage. For reporting years between 2018 and 2023, the HPD program utilized APCD-CDL v2.1, which collected demographic data elements of race, ethnicity, Hispanic indicator, and language. Beginning in 2024, the program transitioned to APCD-CDL v3.0.1, which added gender identity and sexual orientation. Please refer to the APCD-CDL v3.0.1 for more details and accepted values for each element. You may also refer to the [HPD Data Completeness Fact Sheet](#), which provides completeness rates for key fields in each data file, including the member eligibility file.

7. Will you have Kaiser Permanente encounters?

Kaiser Permanente is one of HPD's mandatory submitters that provides all file types to the HPD, including medical claims and encounters.

8. How many years does the data cover?

The HPD program has collected historical data beginning July 29, 2017, and collects data monthly. Public reports and data available through the data access and release program are available for data years between January 1, 2018, and December 2023. HCAI will update these data files on an annual basis for each full year of data received.

9. What is the process for matching multiple episodes for a single individual in the system?

There are HCAI-generated identifiers in the HPD dataset that consistently identify the same individual across years and across plans.

Data Access

10. What is the breakdown of requesters by type of organization (e.g., academic institution, research organization, health plan, hospital)?

As of early January 2026, the breakdown of requests that have been approved or are actively under review is University - 24; Research Organization - 11; Government Entity - 6; Other - 3; Supplier - 1; Provider - 1. HCAI maintains a list of pending and approved data requests on the HCAI website under the "Resources" section of the [HPD Data Access and Release](#) webpage.

11. Is it possible to request Medicare Fee-for-Service or Medi-Cal In-Home Supportive Services (IHSS) data?

HPD collects Medicaid AID category within the medical, pharmacy, and dental claims files. An IHSS recipient may appear in the HPD if the IHSS aid code is provided by the Medi-Cal program.

12. Is Medi-Cal data for county specialty mental health services available in the HPD standard limited dataset, and are there special agreements required for users to access this data?

The HPD Standard Limited Dataset (SLD) includes Medi-Cal data provided by the Department of Health Care Services (DHCS) if DHCS includes county specialty mental health services in its submissions, that data will be part of the SLD. However, identifying county specialty mental health services within the SLD can be challenging because the source system code (which distinguishes these services) is not included.

Access to Medi-Cal data is not automatic. Users must specifically request Medi-Cal data as part of their HPD application. These requests must demonstrate that the proposed use supports Medi-Cal program goals and are subject to review and approval by the DHCS Data Review Committee and Chief Data Officer.

13. For researchers using remote data access (the secure data enclave), what is the process if preliminary analyses suggest a need to modify the original research question or request different data elements?

If an approved use of the non-public HPD data files evolves to include a scope of data use not covered by the initial approved request, there is a process to amend the use case. This is referred to as a supplemental data request.

14. For early-stage health tech companies, what is the most practical pathway to access HPD data for product development and market analysis?

The first place to look for data is the California Health and Human Services (CalHHS) Open Data Portal. These are publicly available data and do not require HCAI review and approval for use. If this does not provide the level of data needed, then requesting access to the confidential data files is the next option. Any individual or organization is eligible to request confidential HPD Data, including but not limited to state agencies, non-profit research institutions, non-profit educational institutions, hospitals, physician organizations, labor unions, self-insured employer plans, and consumer organizations. Submitting a request does not guarantee approval for the use, and some uses, such as potential commercial and marketing uses, will have more extensive review procedures. The request must contribute to fulfilling the goals of the HPD program. To start your application to access HPD data, please register on the HCAI Data Request Portal. Then, please submit a detailed application, answering the questions to the best of your ability and attaching necessary documentation. HCAI will then contact you for the next steps.

15. Are the chronic conditions based on aggregation of ICD-10-CM codes? If so, what aggregation system is used? Is the Clinical Classification Software Refined (CCSR) system used for aggregation?

The chronic conditions are based on aggregation of ICD -10 CM codes and Chronic Conditions Data Warehouse guidelines used to flag the chronic conditions.

16. What are some tips from HPD staff for researchers who are interested in analyzing HPD data?

Review the information available on HCAI's website and the CalHHS Open Data Portal. This includes reviewing our existing public-facing data products and the corresponding technical notes, the data documentation for what is available in the HPD database, being aware of changes in the data from one data year to the next, reading and being aware of the data quality and completeness for those measures you plan to use in your analysis.

17. Will the process for HPD access approval be the same for CalHHS Departments/other State Partners as for external researchers?

When an application is approved, the final steps to complete before accessing data are to have all users sign their corresponding Data Use Agreement, for payment to be received and deposited, and the project space configured for requests accessing data through the Secure Data Enclave. Custom-created datasets also require additional time. The time needed varies depending on the number of users, the method of payment, and the scope of the project space being created. Some lessons learned are to limit any changes to the request, users of the data, project space, or data files requested following approval; and to pay via credit card and to identify any payment-related requirements of your organization early in the request process. A typical project gains access to the data within 30 days of receiving approval.

18. The Justification Grid lists "Denied Claim Flag" and "Denied Claim Line Indicator Code" for Research Identifiable or Custom Limited datasets. Does the data provide specific reasons for denial?

Fully denied claims, in which all service lines have been denied, are not submitted to the HPD. When a claim is partially denied, the entire claim is still submitted to the HPD, and service lines that were denied are indicated by the Denied Claim Indicator field with a value of "1" for denied. A Claim Adjustment Reason Code is collected, which follows the reason code that is maintained by ANSI ASC X12.

19. How long after submitting the application can we expect to receive a review and subsequently access the data?

HCAI's initial review of the application is typically completed within two weeks of the application fee being paid.

Data Request Process

20. Could you please provide the timeline for receiving the results from the price reduction request? Do we need to submit a data request to initiate this process?

Upon receipt of a data request application and the paid application fee, HCAI completes an initial review. HCAI's best practice experience to date is to ask the requester to respond to questions and revisions needed that are identified by the initial review. Doing so provides a more accurate level of detail in considering the price reduction request. A decision regarding the price reduction request typically follows this initial process but can vary due to the nature of the data request. If the data request application is modified following the initial price reduction decision, the price reduction is revisited for final decision-making.

Data File Specifications

21. Are providers and health plans identifiable?

The provider file within the HPD includes information such as the provider's name, national provider identification (NPI), taxonomy codes, and location. The provider file is a roster of providers included within the member eligibility, medical, pharmacy, and dental claims files. Health plan names, along with an HPD assigned payer code and submitter code, are included within the HPD system for all file types collected.

22. How does HPD track unique individuals over time and across different payers (i.e., if a patient changes insurance)?

There are HCAI-generated identifiers in the HPD dataset that consistently identify the same individual across years and across plans.

23. Is it possible to use the Provider File to analyze provider activity, such as tracking participation in different health plans, relative volume of visits by plan, or types of services by plan?

The Provider File is composed of information, including but not limited to provider identifiers, provider names, specialty codes, and practice locations for all providers included in the Member Eligibility, Medical, Pharmacy, and Dental Claims Files. Therefore, the Provider File by itself would not be able to analyze these topics. This type of analysis may be conducted by utilizing the claims files. The claims files include the providers' NPI information associated with the claim lines.

24. Does the provider file include an HCAI identifier or Medicare ID?

The Provider File includes a Provider Medicare Provider ID. For data elements collected as part of the Provider File, please refer to the APCD-CDL available on the [HPD Resources](#) page. For data elements available as part of the data access and release program, please refer to the [HPD Data Access and Release](#) page.

Pricing

25. Will a limited HPD dataset be made available to local health jurisdictions free of cost? What discounts are available for non-profit organizations?

HCAI worked with the HPD stakeholder advisory committee to develop a fair, competitive pricing policy for HPD, which includes the option for requesting a price reduction, with good cause and justification. For example, with sufficient justification, price reductions may be available for entities that:

1. Are consumer organizations, students or academic fellows, government organizations, or data submitters to HPD,
2. Are working on behalf of a non-profit organization that has a demonstrable financial hardship,
3. Are working on projects in high-priority areas, or will lead to innovations that will benefit the public at large, including:
 - Health equity, or
 - Health Workforce, or
 - Affordability, or
 - Produces open-sourced code available to be shared with HCAI, or
 - Results help to evaluate or improve HPD data quality.

These discounts can be partial or full. Please consult with a data request analyst for more information.

Use Cases

26. What is the recommended way to use HPD to quantify and measure access gaps and disparities over time (e.g., by county/ZIP, rural vs urban, socioeconomic proxies, language, etc.)?

To determine whether HPD would be helpful for your specific use case, HCAI recommends reviewing existing products available on HCAI's website, such as Snapshot, Measures, and Services. These products may answer or begin to answer your questions. HCAI also recommends looking at the supportive documentation for products available on our website and the data available on the Open Data Portal. If your use case is not answered by existing products, the next option would be to submit an HPD data request and work with HCAI analysts to examine the detailed use case and linkage possibilities.

27. Is it possible to use HPD data that are publicly released (and/or available for researchers) to evaluate demographic disparities in access to specific services, such as clinical genetic/genomic testing?

Existing products available on HCAI's website – especially Services and Measures – can be used to answer some questions about disparities in access to services. For a specific service not covered in these publicly available data products, such as clinical genetic/genomic testing, users may be able to apply for HPD data and use procedure codes to identify any services for which providers submitted a claim to an HPD data submitter. HCAI recommends submitting an HPD data request and working with HCAI staff to determine whether HPD data would support the analysis of the specific access disparity in question.

28. Can academic investigators collaborate with HCAI staff on shared research interests? Additionally, are there funding opportunities to support student training or investigator research activities?

HCAI does not directly fund academic investigations using the HPD Program Data. There are external funding opportunities, most notably a funding opportunity from the California Health Care Foundation (CHCF). CHCF recently released the [Request for Proposals: 2026 California Healthcare Payments Database Affordability Research Fund](#) for their second year of funding opportunity. HCAI's primary focus in these early stages of the HPD Program is completing public reporting priorities identified for the program and making data available to external users of the HPD Program data. As HCAI's public reporting portfolio matures, there may be more opportunities for formal collaboration between academic investigators and HCAI staff. HCAI staff are available to provide technical assistance to academic investigators.

29. What guardrails and review standards apply when using HPD data to train or validate predictive models (e.g., cost-risk, adherence, care pathway optimization) to prevent re-identification?

When reviewing and approving requests to use the confidential data files, HCAI uses the principle of minimum data necessary for the use case. Some use cases can be accomplished without the use of direct or indirect personal identifiers, and HCAI will work with the requester to achieve the goals of their use case without such identifiers. In the event that identifiers are needed, minimum data necessary will still apply. HCAI also requires that any presentation of the data is de-identified according to the CalHHS de-identification guidelines. HCAI also guides requests towards accessing the data in the Secure Data Enclave, where data and information cannot be downloaded and/or taken out of the Secure Data Enclave without HCAI's review and approval; included in this approval is confirming adherence to the de-identification guidelines. Lastly, training models or validating models via the use of an AI will be limited if not fully restricted.

30. Will this data be linked with the California Cancer Registry data?

An organization interested in linking the HPD with California Cancer Registry data can submit a data request application to HCAI for receipt of the HPD for this purpose; a separate request to the California Department of Public Health (CDPH) will be needed for the California Cancer Registry data. Upon approval from both HCAI and CDPH, the linkage can be completed by the requesting organization.

Enclave

31. When would Medicare FFS claims be available in the Enclave?

Medicare Advantage data is collected from managed care plans operating in California, whereas Medicare Fee-for-Service data is obtained from Centers for Medicare & Medicaid Services (CMS) through a data use agreement between the federal government and the state. The terms of this data use agreement state that Medicare FFS data are only authorized for use by the California State Government. A California State Government request for Medicare FFS data that receives approval will get access to the data following receipt of all signed Data Use Agreements, final payment for the price of the project, and upon configuration of the project within the secure Data Enclave.