Ask the Analyst Webinar Series
September 27, 2023 – Q&A Summary

Below is a summary of the questions and answers discussed during HCAI's Ask the Analyst Webinar Series held on September 27, 2023, including questions received during the event for which there was not time to answer during the webinar. HCAI looks forward to seeing you at the next event! Subscribe to HCAI's mailing lists to be notified of future public events.

The Health Care Payments Database (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 webinar included an overview of the program and virtual tours of the first two HPD data product releases, the HPD Snapshot and the HPD Measures Report. HCAI's panel of subject matter experts responded to questions from the audience before and during the event.

Any additional questions or feedback may be sent to dataandreports@hcai.ca.gov. Please include "Ask the Analyst" in the subject line.

HPD Snapshot

On the Medical Procedures tab, could you explain the "Type of Setting" filter?

Type of Setting categories are assigned using bill type codes, place of service codes, procedure codes, and revenue codes on claims and encounter records. Examples include inpatient, outpatient, pharmacy, provider, home health, and hospice.

When a patient receives multiple services in a hospital, what will the data show for "Type of Setting"?

An individual admitted to a hospital for surgery will generate multiple records for services in different settings. For example, the facility fees associated with the inpatient stay will be classified as "inpatient" setting. The surgeon and anesthesiologist will submit claims (or encounters) classified as the "provider" setting.

The procedure, such as a hip replacement, is counted only once (usually under "provider" setting), but associated services such as physical therapy or lab work may appear multiple times ("inpatient" setting).

On the Medical Procedures tab, can you paste a list of procedures into the "Search by Procedure Category" field?

The search field on the visualization does not have that capability. However, the underlying data is available on the Open Data Portal at the level of the procedure code to allow for analysis of individual services.

Can the "Search by Procedure Category" include a drop down or filters in the future?

The list of procedure category topics is extensive and would make a drop-down menu or a filter a lengthy selection. HCAI is working on the capacity of search features on HCAI webpages and visualizations with the goal of greater responsiveness to a wider range of key words.

HPD Measures Report

Los Angeles County has around 10 million residents, about one-quarter of California's population. Currently, the smallest reporting unit for LA data is two Covered California regions. Is there any plan to report on LA County at a more granular level?

Los Angeles, in particular, has been of interest for additional sub-geographies. The next filter HCAI is assessing and considering for LA County is the Service Provider Area, or SPA. However, as has been a theme in the early stages of data use, data completeness and accuracy are key. HCAI is evaluating the quality and completeness of data fields such as patient billing address to identify any anomalies that could affect analysis and reporting.

HCAI will include a review of data completeness in a 2024 report to the California Legislature.

Why did HCAI choose the health conditions used in the visualizations? And why are some conditions not listed for 2018?

Health conditions are identified based on diagnosis codes and defined using the criteria outlined by the Centers for Medicare & Medicaid Services (CMS) in the Chronic Conditions Data Warehouse. The 23 health conditions were chosen because they are prevalent in California's population, affect individuals at different points in the life span, and affect different body systems. The conditions include both chronic conditions, such as diabetes, and health events, such as stroke. The CMS methodology requires two years of data to accurately estimate the prevalence of chronic conditions. Because HCAI does not have data for 2017, a two-year time series for some conditions for 2018 was not possible.

What is the difference between "emergency department visits" and "potentially avoidable visits" and why is that important?

"Potentially Avoidable Emergency Department (ED) Visits" are a subset of all emergency department visits. An ED visit is "potentially avoidable" when the reason for the visit could have been effectively addressed through primary or urgent care; examples include conjunctivitis, ear infections, and upper respiratory issues. This measure can indicate a lack of access to primary care. It can also highlight potential cost savings since the ED is a more costly care setting than primary care.

How can I identify utilization and chronic condition prevalence for "dual eligibles" – individuals with both Medicare and Medi-Cal?

The HPD Measures Report does not currently allow users to identify dual eligibles. The visualization can only show one "payer type" for each member, so those individuals appear as Medicare or Medi-Cal but not both. Reporting on the dual eligible population is an area of interest for future public reporting.

Does HCAI plan to add payment data, such as "allowed amount", to the HPD Measures Report?

HCAI has plans to incorporate additional data, including payment data, into both existing and forthcoming public data products. The first public report that incorporates payment data will be on prescription drugs, targeted for late 2023; we expect to build on that report in future iterations. However, given the prevalence of capitation and other alternative payment arrangements in California, total cost of care reporting will need to wait until non-claims data is available. HCAI is currently targeting Q3 2025 to begin collecting non-claims data.

General HPD Program

What are some of the next things that you're planning for these products that were released? What might be some of the next data coming soon?

One of the next steps is to address the Medicare fee-for-service data that was not available to include for 2021. HCAI will add these data when these products are refreshed and has developed a plan to keep annual refresh of data as up-to-date as possible, despite delays in receiving the Medicare fee-for-service data from the federal government. HCAI has also received requests from users for additional filters to the reports, such as the ability to filter by a specific payer, filters for additional patient characteristic data like additional geographies for viewing the data. Adding these features will require HCAI to ensure the underlying data quality and maturity of the database is sufficient to support them. HCAI will continue to solicit feedback from its stakeholder, including the HPD Advisory Committee, about what analyses are best supported by the data.

What types of information does HCAI get from claims data versus encounter data? What are the similarities and differences between these two?

Both claims and encounter data capture patient demographic data (patient information, such as date of birth, address, etc.), provider specific data, service code, and health plan information. The difference between the two is that encounter data does not include payment amount, as claims data does. Encounter data serves as a record of rendered services under a capitation-based payment model rather than for billing.

What does it mean that HPD reports the data as submitted?

The HPD reports the data submitted by payers to HCAI. While HPD will enhance the data by adding groupers, calculated fields, and flags that aid analysis, HPD will not change the data submitted by payers. Therefore, the HPD reports and datasets may include errors and omissions that were contained in the original submitted data. HPD works regularly with payers to improve data quality and completeness over time.

Has HCAI considered adding filters to the Measures Report that would allow me to view differences in rates of chronic disease or utilization by race/ethnicity, sexual orientation/gender identity, income, or other relevant factors?

HCAI is actively evaluating how HPD data can be used to report on and understand health disparities and improve health equity, in collaboration with the HPD Advisory Committee.

Data quality and completeness is an important factor to being able to appropriately analyze patient populations. HPD data is reported by HCAI "as submitted," which means that while HPD will enhance the data by adding groupers, calculated fields, and flags that aid analysis, HPD will not change the data submitted by payers. HPD works regularly with payers to improve data quality and completeness over time. HCAI will include a thorough review of data completeness in a 2024 report to the California Legislature.

Why do you only report by primary payer?

Assigning an individual to one payer type prevents counting someone twice or using their data multiple times when calculating a measure. There are many different combinations of coverage in HPD; by focusing only on the primary payer in the HPD public reports, we eliminate the need for subjective decisions about when and how to account for secondary coverage. Some topics for potential future reporting and analysis, such as analysis of the "dual eligible" population (those covered by both Medicare and Medi-Cal), will need to account for more than one source of coverage as a central component. HCAI will be evaluating how to support such analyses in the future.

What is the impact of the missing Medicare fee-for-service data? How will the data change once the Medicare fee-for-service data is added?

The missing Medicare fee-for-service (FFS) data results in an undercount of Medicare and an overcount of Medi-Cal for reporting year 2021. The overcount of Medi-Cal enrollment occurs because Medicare is the primary insurer for individuals with both Medicare and Medi-Cal eligibility ("dual eligibles"), but dual-eligibles with missing Medicare FFS data in 2021 will appear in the data as Medi-Cal only. In addition to the under- and overcounts, trends will be impacted, rates can be overestimated or underestimated, and certain populations – such as older age groups – will be impacted usually as an undercount. In these products the data is only missing from 2021, and so it is the 2021 data that should be used cautiously. HCAI will add these data when these products are refreshed and has developed a plan to keep annual refresh of data as up-to-date as possible, despite delays in receiving the Medicare fee-for-service data from the federal government.

Why does it say "Suppressed" on some of the graphs?

All filters and grouping dimensions are helpful to create subgroups and answer specific questions, but dissecting the data so many times can lead to very small groups. Even with aggregate or summarized data, we want to reduce the risk of individuals being identified by the characteristics of these groups. As with any HCAI report where data is released publicly, HPD follows the CalHHS Data De-Identification Guidelines. Given the complexity and size of the underlying data, one of the guidelines we use is to remove data that has a group size of less than 11 and you'll see "Suppressed" instead of the specific value, in order to protect individual privacy.

Can you download this information into a CSV or an Excel file?

The deidentified underlying data for the reports that we've published is available in a CSV file format. HCAI has been a participant of the open data movement for nearly a decade, demonstrating a strong commitment to making data available in accessible, machine-readable formats that can be used in software applications and other tools. The underlying data for both HPD Snapshot and HPD Measures is available on the Open Data Portal in CSV file format. Additionally, there is an API service that also allows API connectivity to that dataset accessed by clicking the dataset name and clicking the API button.

If you would like to see the documentation for how to code against the APIs that are enabled for other datasets, you can visit this link to the Open Data Portal API guide.

Is the information always shared in the aggregate or can it be identified to specific data submitters?

HCAI produces aggregated public reports based on feedback from HCAI stakeholders, including the HPD Advisory Committee. HCAI anticipates continuing to advance the accessibility and usefulness of HPD public reports as the database becomes more comprehensive and complete and as HCAI builds its capacity over

time. HCAI is also in the process of developing a data release program that will allow eligible requestors to perform their own analyses on record-level HPD data. The HPD Data Release Committee is in the process of advising HCAI on policies and procedures for the data release program.

For more information, visit the HPD Data Release Committee page.

I am interested in the potential to link to the California Cancer Registry data. Is this something that will be available within the HPD database?

The HPD Data Release Program intends to support data linkage as part of its offering for allowing controlled access to record-level HPD data with different file types with varying levels of direct and indirect identifiers that can support data linkage.

The HPD Data Release Committee is currently in the process of advising HCAI on the development of the data release procedures. For more information, visit the HPD Data Release Committee page.

What is the timeline to review completeness and quality of additional demographic data?

HCAI is evaluating the completeness and analytic methods for studying patient demographics, such as for race and ethnicity, so that these important dimensions of health equity can be appropriately examined in future HPD reporting. HCAI will include a review of data completeness in a 2024 report to the California Legislature.

What is the lag time between data submission to HCAI and its inclusion into the database?

Data is received from commercial submitters on a monthly schedule and is due by the first business day of the second month following the month data is reported. For example, March data would be due to the HPD System by the first business day of May. An internal analytic dataset is produced on a quarterly basis from which standardized datasets are created that can comprise complete years of data. Public reports are intended to be refreshed on an annual basis with complete years of data.

Other HCAI Data Programs

I can see a certain year's hospital annual utilization report and can disaggregate by either age OR primary diagnosis but can't figure out how to do the crosstab for that; is there a way for me to do it on my own without a custom data request?

The Hospital Annual Utilization Report does not disaggregate to that level, however, the "Inpatient Hospitalizations and Emergency Department Visits for Patients with a

Behavioral Health Diagnosis in California: Patient Demographics" report includes the percentages for mental health issues versus other diagnoses including substance use disorders, co-occurring disorders, and a combined count of all diagnoses. These data are then disaggregated by age band and setting (inpatient hospitalizations vs. emergency department visits). This report has data from 2020 and 2021, and the 2022 update is expected to be released in the Fall of 2023. Additionally, all data available on the Open Data Portal have the counts, totals, and percents. Further breakdowns of these data such as specific mental health diagnosis categories by age group require a custom data request.