

# An Overview of Claims Data and How CHBRP Estimates Utilization and Unit Cost

OCTOBER 2021



California Health Benefits Review Program (CHBRP)  
University of California, Berkeley

At the request of the legislature, the California Health Benefits Review Program (CHBRP) provides independent evidence-based analyses of proposed health insurance benefit related legislation that would impact state-regulated health plans and policies. To estimate how new or altered benefit coverage might impact utilization and additional expenditures, CHBRP develops a baseline estimate of utilization and unit cost of the specific test, treatment, or service under consideration.

Claims data, also known as billing data or administrative data, is one important tool that CHBRP uses to estimate utilization and unit cost. While analyzing proposed health insurance benefit related legislation, CHBRP works closely with its contracted actuaries, who have expertise in leveraging claims data to support CHBRP's analytic efforts. This explainer provides an overview of claims data, including how the data is collected, what information may be available within the datasets, and how the data are used to inform CHBRP's analyses.

## What is a Claim?

Claims are submitted by healthcare professionals and facilities to enrollees' health insurance plans or policies as a means of receiving reimbursements for tests, treatments, or services (NLM, 2021). Each claim contains procedure and diagnosis codes that are needed for reimbursement, which include information about the services provided and the patient diagnoses requiring treatment.

Healthcare professionals and facilities are reimbursed differently depending on the setting. The four major categories are: inpatient, outpatient, professional, and pharmacy.

- **Inpatient facility claims** are reimbursed using diagnosis-related groups (DRGs), per diems, or as a percentage of billed charges.<sup>1</sup>

<sup>1</sup> Inpatient facility claims are usually reimbursed a set amount. For example, if maternity care is reimbursed on a DRG basis, a health benefit mandate that requires additional coverage of a test, treatment, or service related to the maternity admission would likely not result in any net new reimbursement from a health plan or policy.

- **Outpatient facility claims**, which include services provided in outpatient clinics, emergency departments, ambulatory surgery centers, and outpatient diagnostic and testing departments, only incurs facility charges and are generally reimbursed as case rates.<sup>2</sup>
- **Professional reimbursement** for physicians and other billable professionals usually occurs on a fee schedule based on Current Procedural Terminology (CPT) codes. Since fee schedules vary between plans, policies, and types of professionals, claims data in this setting can strongly reflect the distribution of reimbursement levels seen throughout California. It is also very common in California for at least some of the professional services to be reimbursed in a capitation arrangement rather than as fee-for-service. This distinction is discussed in the “Strengths and Limitations of Claims Data” section.

Inpatient and outpatient services usually have a facility component and a professional component that are billed separately. Professional and facility claims differ in that professional claims include services provided by physicians and other healthcare professionals, whereas facility claims involve services that represent what resources have been consumed (equipment and supplies, laboratory and radiology services, etc.).

- **Pharmacy reimbursement** is usually a discount off of average wholesale price (AWP) for brand name prescription drugs or a fee schedule for generic drugs. Pharmacy claims are almost always submitted electronically, and there is generally more consistency and greater volume in pharmacy claims data.<sup>3</sup>

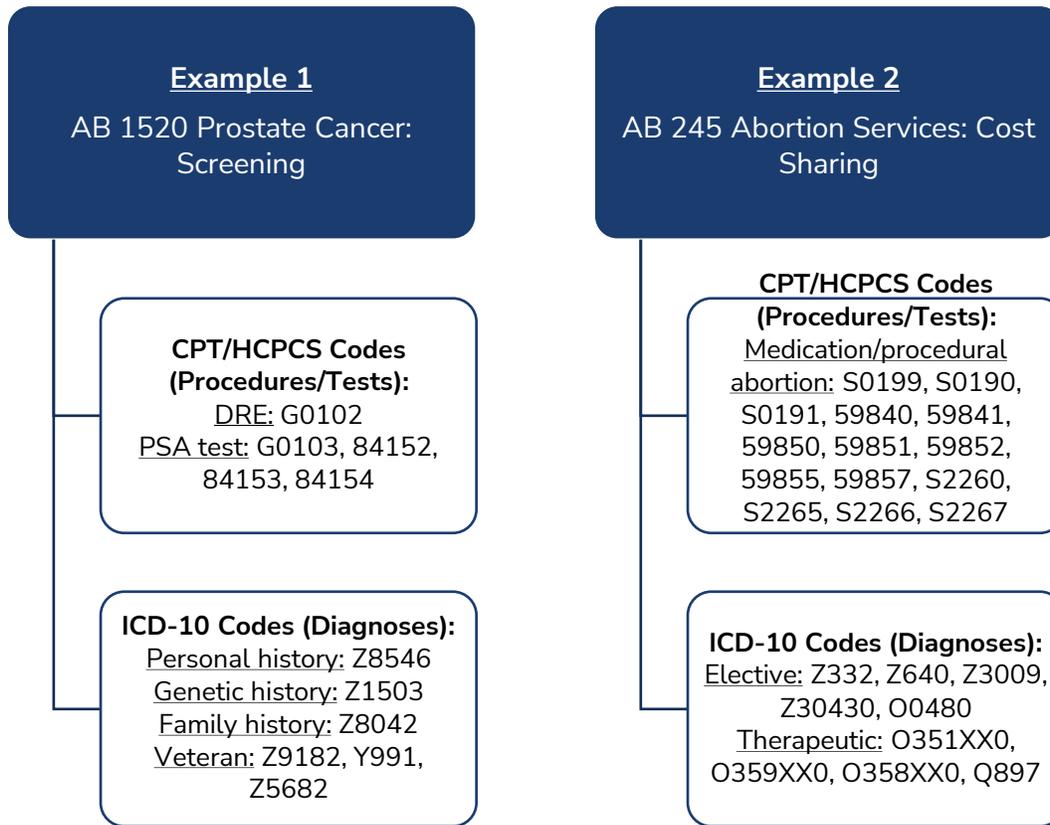
The Centers of Medicare and Medicaid Services established billing code systems under the Health Insurance Portability and Accountability Act, and these systems outline specific revenue, procedure, diagnosis, and drug codes that can be entered into claims forms (ASHA, 2021). The information on claims is recorded through two main types of codes: Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) codes and International Classification of Diseases (ICD-10) codes. CHBRP uses CPT/HCPCS codes to identify enrollees with a specific procedure, service, device, or treatment, and ICD-10 codes to identify enrollees with a specific diagnosis (see Figure 1). While ICD-10 codes do not represent expenditures, charges, or services provided, the diagnostic information captured in ICD-10 codes informs corresponding CPT/HCPCS codes (and vice versa), which do translate directly into medical expenditures and reimbursements (CDC, 2015).

A third type of code, National Drug Code (NDC), identifies “prescription drugs, over-the-counter (OTC) drugs, and insulin products that have been manufactured, prepared, propagated, compounded, or processed by registered establishments for commercial distribution” (FDA, 2017). NDCs are included in claims along with applicable CPT/HCPCS codes when medications or vaccines are billed (AAP, 2016). CHBRP uses NDCs to estimate baseline utilization and unit cost of specific drugs. An example of CHBRP’s use of NDCs is discussed in the “Claims Data in CHBRP Analyses – An Example from the Analysis of Assembly Bill (AB) 97 (2021)” section.

Figure 1 demonstrates how CHBRP uses CPT/HCPCS codes to identify enrollees with a specific procedure, service, device, or treatment and ICD-10 codes to identify enrollees with a specific diagnosis. The two examples present how CHBRP identified enrollees receiving prostate cancer screening (Example 1) or abortion services (Example 2) as part of two analyses completed in 2021. First, claims were subset to only include enrollees with specific CPT/HCPCS codes. Then, enrollees with “high risk” statuses for prostate cancer (Example 1) or with diagnoses for abortions (Example 2) were identified by searching through enrollees’ claims for specific ICD-10 codes.

<sup>2</sup> A case rate is a flat fee reimbursed for services provided for a pre-defined episode of care.

<sup>3</sup> An important consideration when conducting analyses related to emerging pharmaceutical treatments is that pharmacy practice patterns change more frequently than the other settings. Claims data, however, provides a distribution that reflects changes in pharmacy practice patterns and that captures the mix of different pricing guarantees for different dispensing channels.



Source: California Health Benefits Review Program, 2021.

Note: AB 1520 and AB 245 were analyzed by CHBRP in 2021. More information about specific analyses is available at CHBRP’s website under “Completed Analyses”.

Key: CPT/HCPCS = Current Procedural Terminology/Healthcare Common Procedure Coding System; ICD-10 = International Classification of Diseases, Tenth Revision.

## Unit Cost and Allowed Cost

CHBRP often refers to “unit cost” in its analyses. This nearly always means “allowed cost”, which is the maximum amount a health plan or insurer will reimburse professionals and facilities for a covered test, treatment, or service used by plan or policy enrollees (OFM, 2015).

When analyzing a proposed benefit mandate, it is often most useful to review the allowed cost because it serves as a common thread across different health plan or policy designs which may have different levels of enrollee cost sharing. Negotiated prices between health plans or insurers and professionals or facilities are generally confidential, but claims data reflect real negotiated prices that correspond to allowed costs. Moreover, because California has clear regional separations in health care costs, claims databases are beneficial for building distributions to provide an average of allowed costs that represents existing differences (CHCF, 2021).

## Strengths and Limitations of Claims Data

Claims data can be helpful when answering questions about unit cost and utilization, but there are some limitations.<sup>4</sup> Claims databases are useful because they contain data that are longitudinal, multi-sourced, robust, and easily accessible. They also include procedure, diagnosis, and drug codes that indicate the services that have been provided and the resources that have been consumed, which allows CHBRP to estimate utilization for tests, treatments, or services and unit costs reimbursed by plans and insurers.

Claims data is often the best source of available data to estimate utilization and unit cost, but it can be limited by the quality of the information collected (Rowley, 2014). California is unique with its relatively high levels of capitated reimbursements, which can make it harder to obtain a large amount of comprehensive claims data as a result of fewer incentives to submit detailed encounter data (claims that measure quality but are not associated with reimbursements). As an increasing number of health plans become risk adjusted with an emphasis on quality measures, historical fee-for-service claims data can remain useful to help establish capitation rates for providers.

Claims data provides insights into professional practice patterns and different reimbursement settings. Pharmacy practice patterns change most frequently out of the four major categories of claims, but claims data can capture changes in pharmacy practice patterns as well as mixtures of various pricing guarantees. Claims data also captures the mix of how professionals and facilities in California are reimbursed.

Claims data is ultimately a sample of utilization and allowed cost for the California population, which may or may not be representative. Therefore, CHBRP's cost team and contracted actuaries work with the public health team and scale prevalence numbers to more accurately reflect California's population.

## Claims Data in CHBRP Analyses – An Example from the Analysis of AB 97 (2021)

In 2021, CHBRP analyzed AB 97, a proposed health insurance mandate that would prohibit the application of deductibles for insulin prescriptions.<sup>5</sup> In this analysis, CHBRP used relevant codes in claims data in conjunction with other datasets to estimate baseline and postmandate utilization and unit cost of insulin.

National Drug Codes (NDCs) for insulin were identified and used to develop estimates of baseline utilization and allowed cost for insulin users. By analyzing claims data to identify the number of enrollees using insulin at baseline (118,014 enrollees) – stratifying enrollees whose insulin was subject to a deductible at baseline and those whose insulin was not – CHBRP estimated the same baseline utilization rate (0.85 insulin prescriptions per month per user) for all enrollees with health insurance subject to the bill. Claims data was also used to calculate the allowed cost of insulin; CHBRP estimated an average cost of \$491 per insulin prescription per month.

More detailed information on data sources and methods, as well as general caveats and assumptions applicable to CHBRP's cost impacts analyses, are available at CHBRP's website.<sup>6</sup>

<sup>4</sup> Occasionally, claims data is insufficient for CHBRP's analyses, in which case information about utilization and unit cost may be extrapolated from other sources (published studies, gray literature, content expert estimation, etc.).

<sup>5</sup> See CHBRP's report *Analysis of California Assembly Bill 97: Insulin Affordability*. Available at: <http://analyses.chbrp.com/document/view.php?id=1557>.

<sup>6</sup> See CHBRP's cost impact analysis methodology documents, available at: [https://chbrp.com/analysis\\_methodology/cost\\_impact\\_analysis.php](https://chbrp.com/analysis_methodology/cost_impact_analysis.php); in particular, see *2021 Cost Impact Analyses: Data Sources, Caveats, and Assumptions*.

## Conclusion

Claims data, also commonly referred to as billing data or administrative data, is a large-scale record of professional and facility reimbursements as well as capitated encounters, which are not separately reimbursed (NLM, 2021). The information captured in claims data serves as a valuable tool for understanding and analyzing proposed health insurance benefit mandates, particularly when estimating baseline utilization and unit cost of relevant tests, treatments, or services. When CHBRP refers to unit cost in its analyses, CHBRP is nearly always referring to “allowed cost”, which is the maximum amount that health plans and insurers reimburse for a covered test, treatment, or service. Claims data in aggregate provide distributions of allowed cost that reflect the varying levels and settings of professional and facility reimbursement throughout the state. Although valuable for understanding utilization and unit cost, there are data limitations depending on the quality and completeness of the submitted data. Distributions and patterns reflected in claims data may be highly reliable on their own, but these data can be more helpful when considered in conjunction with other data sources and public health findings. CHBRP’s emphasis on collaboration between the cost team and public health team demonstrate how claims data can effectively contribute to analyses of proposed health benefit mandates.

## REFERENCES

American Speech-Language-Hearing Association (ASHA). Introduction to Billing Code Systems. Available at: [https://www.asha.org/practice/reimbursement/coding/code\\_intro/](https://www.asha.org/practice/reimbursement/coding/code_intro/). Accessed August 13, 2021.

California Health Care Foundation (CHCF). California Health Care Almanac: Regional Markets; 2021. Available at: <https://www.chcf.org/resource/california-health-care-almanac/almanac-regional-markets/>. Accessed September 9, 2021.

Centers for Disease Control and Prevention (CDC). Interpreting Health Claims Data; 2015. Available at: <https://www.cdc.gov/workplacehealthpromotion/model/healthcare-cost-data/interpreting.html>. Accessed August 13, 2021.

National Library of Medicine (NLM). Claims Data. Available at: [https://www.nlm.nih.gov/nichsr/stats\\_tutorial/section3/mod3\\_data.html#:~:text=Claims%20data%2C%20also%20known%20as,and%20other%20patient%2Dprovider%20communications](https://www.nlm.nih.gov/nichsr/stats_tutorial/section3/mod3_data.html#:~:text=Claims%20data%2C%20also%20known%20as,and%20other%20patient%2Dprovider%20communications). Accessed August 13, 2021.

Rowley R. Claims Data – The Good, The Bad and The Ugly. Health IT Answers; 2014. Available at: <https://www.healthitanswers.net/claims-data-the-good-the-bad-and-the-ugly/>. Accessed August 13, 2021.

U.S. Food & Drug Administration (FDA). National Drug Code Database Background Information; 2017. Available at: <https://www.fda.gov/drugs/development-approval-process-drugs/national-drug-code-database-background-information>. Accessed October 5, 2021.

Washington Office of Financial Management (OFM) Forecasting and Research Division. Definitions for billed, allowed and paid amounts and other payment terms. 2015. Available at: [https://ofm.wa.gov/sites/default/files/public/legacy/healthcare/pricetransparency/pdf/phase\\_one/aper\\_3.pdf](https://ofm.wa.gov/sites/default/files/public/legacy/healthcare/pricetransparency/pdf/phase_one/aper_3.pdf). Accessed August 13, 2021.