Clarify

Bringing Clarity to Healthcare Price Transparency Data

Methodology for ingestion, enrichment, and conversion into easy-to-query rate intelligence

MAY 2024

WHITE PAPER

Table of contents

Executive Summary

After decades of price negotiations between health insurance companies and healthcare providers being cloaked in secrecy, payers and hospitals are now required to publicly share negotiated rates. However, the public rate data that has been released is complex and challenging to work within its raw state. Raw rates data released by payers contains trillions of records, is not consistently formatted, and includes payer-specific gaps in reporting, creating a web of convoluted healthcare pricing in the commercial market. Hospital-reported rates data is beset by lower coverage, even less standardization, and is reported for a smaller set of inpatient services. However, for those with the infrastructure to parse through the noise, there is plenty of value to be uncovered from price transparency data.

Since 2015, the Clarify Atlas Platform® has been managing massive healthcare data sets using big data efficiencies. Our investments in processing power provide the perfect infrastructure to handle the size and scale of price transparency data. As a result, we can provide price transparency intelligence that helps healthcare organizations stay ahead of the curve by strengthening contract negotiations and improving provider network affordability.

We deliver rate intelligence via a powerful no-code query engine and at-a-glance, actionable dashboards in cloud software. Healthcare organizations can see their market pricing position instantly and easily query over 1.5 petabytes (1500+ terabytes) of enriched rates data and generate reports on market prices in seconds. This puts the power in the hands of the end user, reducing the need for a fully staffed analytics team to wade through the vast sea of price transparency data. Aside from the ability to query data without coding, what makes our solution unique is our proprietary data enrichment process. By enriching raw rates with claims data on 300M+ annual lives, the product displays prices that have been billed instead of millions of meaningless prices.

In this paper, we describe the robust methodology that we have developed to convert price transparency data from multiple sources into easy-to-query rate intelligence. We detail our data ingestion, cleaning, and proprietary enrichment process that is first-of-its-kind in the industry. Our enrichment process combines claims data with trusted clinical informatics methodology to map standard codes and service lines, estimate Medicare payments for those services, and calculate rates as a percentage of Medicare payments. It assesses claims volume to determine utilization and remove 'zombie rates,' rates that are reported for providers who never have and never would provide certain services. Clarify adds the critical dimensions of service mix and volume by integrating insights from claims alongside rates data. This also enables the display of the weighted percent of Medicare prices (versus straight average) so users can better inform their understanding of the true economics and value delivered.

With a better understanding of our methods, we hope you can determine if our approach is suitable for your technical and business needs.

> "With Clarify's rate data, our team can now leverage the combined power of Clarify's provider quality performance insights in tandem with rate insights to fully assess
> provider performance and help our members get the higher quality outcomes they deserve at an affordable price."
>
> VP, MARKET AND PROVIDER ANALYTICS
> NATIONAL HEALTH PLAN

The new era of price transparency

As healthcare costs continue to climb in the US, healthcare decision-makers search for strategies to curb those ever-growing costs. After decades of price information cloaked in secrecy, in January 2021, the Healthcare Price Transparency Act went into effect, requiring hospitals to publish their cash pay rates and rates negotiated with health insurance companies. Closely following, the Transparency in Coverage Final Rule required most health plans to share negotiated rates in publicly available files and provide out-of-pocket cost information to covered members. Beginning on July 1, 2022, health insurers were required to share their negotiated rates for all in-network covered services and items as well as their out-of-network allowed amounts and billed charges. By January 1, 2023, cost-sharing estimates for 500 common shoppable services became available in consumer-friendly formats, and by 2024 the remainder of covered items and services must be made available.

Hospital Price Transparency Rule (HPTR)

Hospitals, as of January 1, 2021, are required to publish not only their chargemaster of list prices but also their allowed amounts for each contracted payer network, their cash discount rate, and the minimum and maximum payments for those services. CMS requires these to be published in a machine-readable file (MRF). Hospitals must also publish the most commonly scheduled procedures figures to an interactive dashboard that prospective patients could peruse before choosing a provider. CMS has recommended hospitals report on up to 500 shoppable services to aid consumers. Penalties for hospitals failing to comply will incur a minimum Civil Monetary Penalty (CMP) of $300 a day. Despite these stipulated penalties, both compliance and enforcement of the HPTR have been poor. Not all hospitals, including many regional health systems, have complied with the requirements to make rates available. Current estimates of hospital compliance range from 25 to 70 percent, leading to calls by Congress for more federal oversight and enforcement to increase hospital compliance. In July 2023, CMS proposed enhancements to these regulations, emphasizing increased compliance enforcement and increased standardization of data using CMS templates and MRF accessibility requirements. These changes went into effect in January 2024.

Transparency in Coverage (TiC) Rule

As of July 1, 2022, health plans and self-insured employers were required to disclose their contracted in-network rates with all provider sites in their network, not only for hospitals but also organizations such as physician groups and ambulatory surgical centers.

Like the HPTR, the TiC rule requires payers to publish their pricing data as MRFs containing the following sets of costs for items and services:

1. In-Network Rate File: Rates for all covered items and services between the plan or issuer and in-network providers.

2. Allowed Amount File: Allowed amounts for and billed charges from out-of-network providers.

3. Prescription Drug: Negotiated rates and historical net prices for prescription drugs.

CMS penalties for payers failing to comply were set out in existing law (Public Health Services Act and ERISA), which is $100 a day per impacted individual. Relative to HPTR penalties for hospitals, $100 per insured member per day is a far larger amount, and likewise, compliance by payers has been high. While some gaps are evident in Clarify's review of the payer price transparency data, as discussed below, nearly 100% of all US health insurance plans are partially compliant with TiC regulations.

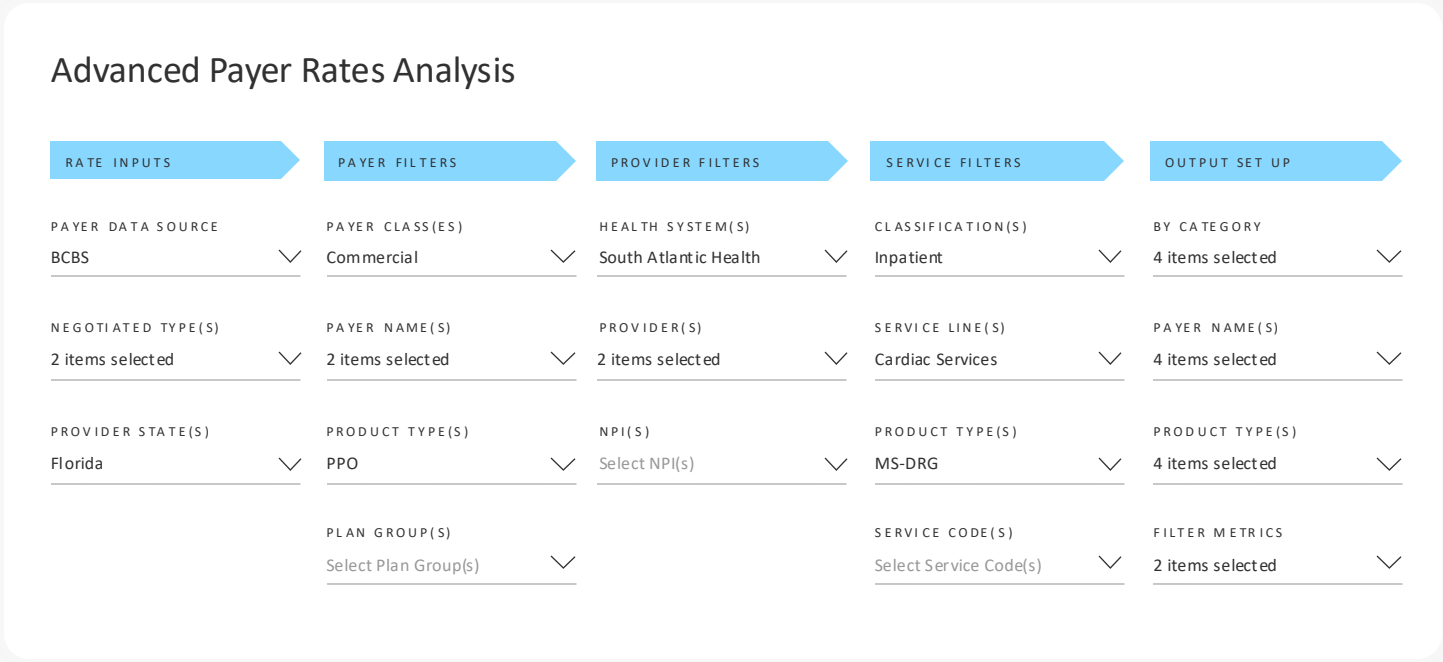
Clarify's price transparency solution

Clarify Rates IQ Suite

Clarify Rates IQ Suite, powered by the Clarify Atlas Platform, is on-demand software for rich intelligence on healthcare rates negotiated between health insurance companies and providers. It offers a competitive edge in contract negotiations and helps build more affordable provider networks.

With 1.5 petabytes of price transparency data that has been enriched with Clarify's claims data, consisting of 300M+ annual lives, users can download hundreds of prices that have been billed (instead of viewing millions of meaningless prices). Additionally, by integrating insights from claims alongside rates data, Clarify adds the critical dimensions of service mix and volume. This also enables the display of the weighted percent of Medicare prices (versus straight average) so users can better inform their understanding of the true economics and value delivered.

With an intuitive user interface and without the need to code, Clarify delivers insights in seconds instead of hours. It can instantly run queries on our massive data set, dynamically filter prices by over 20 categories, benchmark rates to Medicare and regional averages, and even create executive-level dashboards. It can compare inpatient, outpatient, and professional rates to the market, display as a percent of Medicare and drill down at the individual payer, provider, DRG, and CPT code levels.



Clarify Atlas Platform

The Clarify Atlas Platform houses the healthcare industry's largest collection of patient journeys, delivering the most precise and actionable insights to payers, providers, and life sciences companies. It maps over four billion patient journeys across over 300 million lives, leverages best-in-class grouper technology, and drives 20B+ AI-powered predictions, answering healthcare's most complex business questions in an instant. Leveraging patient journeys to enrich price transparency data provides a cleaner data set for unparalleled access to rate intelligence that is truly meaningful.

TRUSTED HEALTHCARE DATA

Atlas brings together 15+ billion government and commercial claim records and Rx data covering 300+million lives with social determinants of health factors and 1.5+ PB of price transparency data in secure, HIPAA-compliant data lakes.

FAST PROCESSING

The tech stack behind our analytics platform can load and process billions of claims in hours and enable the visualization of 15 million episodes of care in a secure, cloud-based SaaS platform in seconds.

TRANSPARENT INSIGHTS

We train thousands of models, cut data in countless ways, and ensure optimal case-mix adjusted predictive values to deliver clear and actionable insights.

A SINGLE PLATFORM TO POWER INSIGHTS FOR ALL

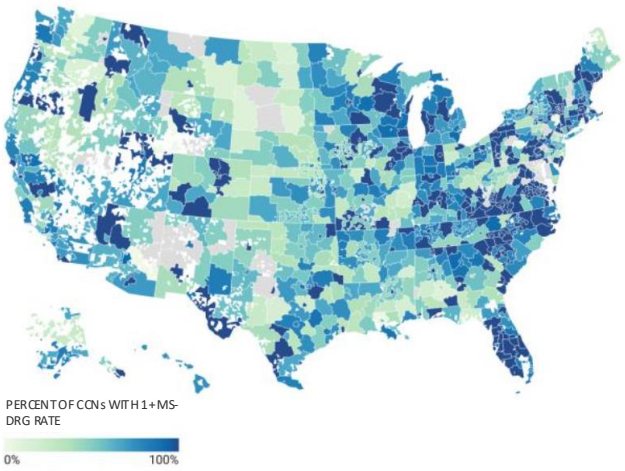
All business applications for payers, providers, and life sciences are powered by a single, modular healthcare analytics platform relying on the same core healthcare data and architecture, data science models, and delivery mechanisms

Addressing data gaps and achieving greater completeness

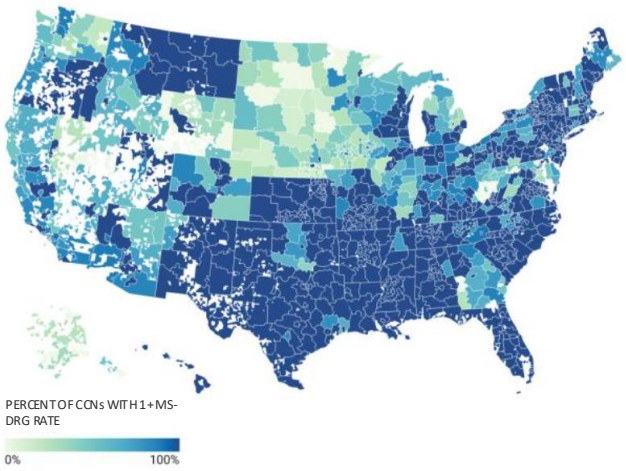
Clarify has processed, analyzed, and compared data from approximately 5,600 hospitals and all large national payers since the inception of each data set to assess quality differences between the data sets. We observe that, as of April 2024, health plans are more compliant with price transparency regulations than hospitals, resulting in higher data coverage across payer-sourced data.

FIGURE 1. PAYER-SOURCED RATE DATA HAVE HIGHER COVERAGE & COMPLIANCE.

Hospitals (identified by CMS CCN IDs) reporting at least one MS-DRG hospital rate:



CCNs with at least one MS-DRG payer rate reported by Aetna/CVS, BCBS, Centene, Cigna, United Healthcare:



Relative to payer reporting, hospital price transparency reporting has often been lower due to a weaker regulatory framework (e.g., no initial requirements for standardization), resulting in worse compliance by hospitals and, therefore, lower coverage across hospital-sourced rates compared to payer-sourced rates. However, we recognize that compliance with price transparency efforts has been mixed for both hospitals and payers. So, in addition to assessing coverage, we have compared the quality of payer and hospital-reported rates data and find significant value in leveraging both datasets together.

- Hospital-sourced rates data is limited to hospital services (e.g., it is missing professional and ambulatory services), whereas over 90% of unique negotiated rates reported by health plans reflect rates negotiated with professionals.

- Notably, CMS also did not require a standard schema for hospital MRF formatting, whereas health plans must follow a defined schema. This has led to a wider variation in hospital reporting format and quality.

FIGURE 2. CLARIFY'S ANALYSIS OF HOSPITAL-SOURCED AND PAYER-SOURCED RATES DATA QUALITY

		HOSPITAL DATA	PAYER DATA
Data Insights	Institutional data	✓	✓
	Professional data		✓
	Employer plan data		✓
	Medicare Advantage and Managed Medicaid plan data	✓	
	Larger reported code & service basket		✓
	Signal on site of service	✓	✓
	Negotiation arrangement information		✓
Data Fidelity	Gaps in coverage	✓	✓
	Standard schema defined by CMS	✓	✓
	Higher refresh cadence		✓
	Higher compliance rate		✓
	Higher penalties for non-compliance		✓

Use of both payer and provider data sources allows cross-validation, gap-filling, and consideration of a broader set of negotiated rates compared to either set alone. Even though payer-sourced rates data often provides higher coverage and quality than provider-sourced data on its own, the Clarify Rates IQ Suite includes hospital-sourced rates in addition to the payer-sourced rates for a more complete dataset.

Data preparation

We have fully extracted, cleaned, and enriched these data for over 65 national and regional payers as of April 2024, including United Healthcare, Cigna, Aetna CVS, Humana, and multiple Blue Cross Blue Shield (BCBS) payers. Clarify has also ingested MRFs from 5,600 hospital facilities.

Acquiring, scraping, and processing these MRFs can be challenging, given their sheer magnitude and complexities created by some payers, making their full file set more difficult to download or process. A high-level overview across twelve payers (including all BCBS payers) is provided below, covering both file counts and coverage in terms of unique service codes and national provider identifiers (NPIs) as of April 2024.

1.5 PB

Since the release of payer MRFs in July 2022, Clarify has ingested over a petabytes of raw, compressed data.

FIGURE 3. PAYER MRF QUALITY STATS (LAST REFRESHED OCTOBER 2023)

ISSUER	# FILES DOWNLOADED	SIZE OF FILES DOWNLOADED	TOTAL UNIQUE MS-DRGs REPORTED	TOTAL UNIQUE HCPS/CPT REPORTED	TOTAL UNIQUE NPIs REPORTED
Aetna	39,081	620 TB	893	22,534	1,284,239
BCBS	20,190	53.7 TB	900	380,091	2,292,364
Centene	78	2.8 MB	683	11,374	919,533
Cigna	84,509	683 GB	898	30,701	1,480,000
Humana	99,000	249 MB	794	22,986	1,144,617
Kaiser	16,040	6 GB	801	34,168	783,921
Medica	137	163 GB	734	18,615	475,600
Medical Mutual	13	20.6 MB	16	2,103	58,475
Moda Health	58	379 GB	893	22,998	1,198,000
Molina	6,547	167 GB	0	14,131	363,245
Optima Health	5	16.45 GB	951	18,050	204,565
Oscar	117	24,8 GB	0	7,237	534,924
PacificSource	289	120 GB	--	54,658	23,009
Providence	6468	3.63 TB	804	28,367	501,421
SelectHealth	354	3.12 GB	--	18,265	41,558
UHC	54,778	59.4 TB	769	16,175	1,815,844
UPMC	4549	14.9 GB	767	12,916	71,000
WellMed	393	6.9 GB	679	15,012	937476

Addressing gaps and missingness

We comprehensively review each payer's MRF during our cleaning and enrichment processes. Through this process, we identify gaps in reporting by some payers. We communicate with payers about these gaps and expect unintended non-compliance with the price transparency regulations to improve over time. Examples of identified gaps or complications in processing from the above set of payers include:

Both Aetna and BCBS MRFs have, for some providers, included uniform rate amounts across tens to hundreds of different MS-DRGs. This duplication is likely a reporting error on the part of the payer or their vendor responsible for creating the MRF.

BCBS rates (compiled from over 40 individual BCBS payers) also contain substantial duplication due to BCBS payers' reciprocity agreements, such as the BlueCard program, and the payers' interpreted compliance with regulations requiring reporting of all reciprocal rates by the specific BCBS payer.

Centene's overall MRF size is observed to be smaller than other national payers, resulting in various gaps in specific rates and service codes.

Cigna institutional rates are typically reported as individual rather than organizational NPIs (over 90%). In addition, Cigna broadly applies a 'CSTM' approach allowed under the transparency in coverage regulations. Payers may use code type CSTM-ALL and the code CSTM-00 as a type of global rate reporting. Per CMS, this code type "represents all possible coding types under the contractual agreement" and the code "represents all possible billing code values", complicating analyses.

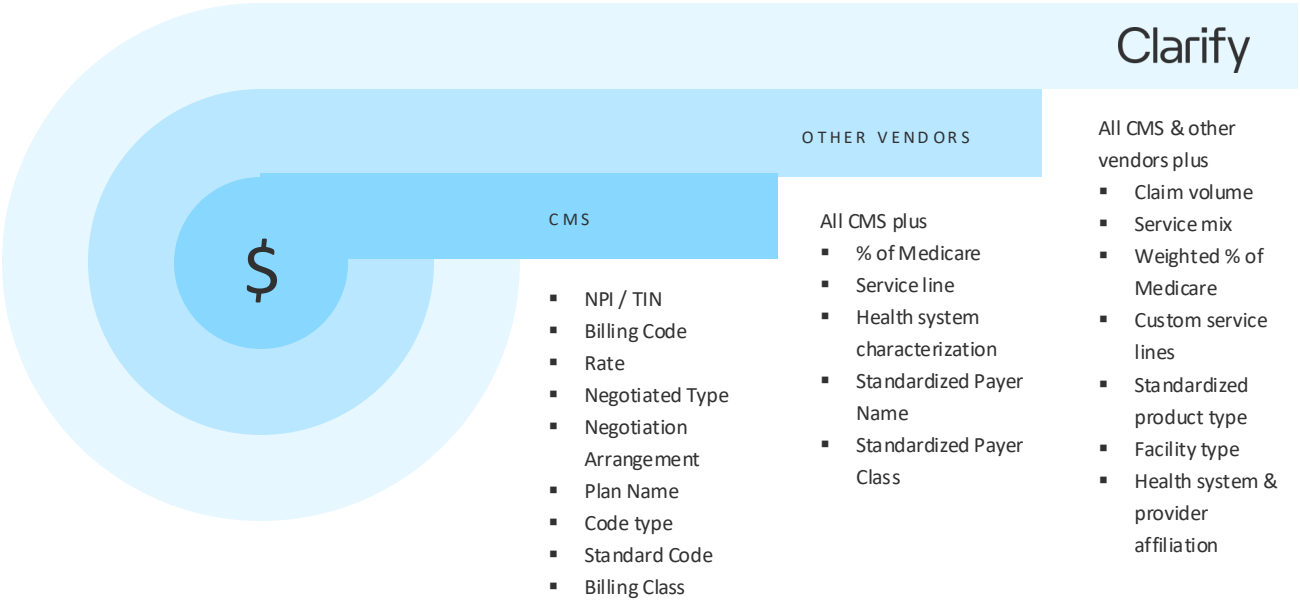
Humana's MRFs are massive, and their posted files total over 400 terabytes. Humana was the only national insurer to initially post their data in a text format (.CSV). While JSON formats have been posted recently, Humana's MRFs remain the largest, which results in slower processing speed.

UHC's rates data underreports institutional outpatient rates, with between 25-50% of unique HCPCS for outpatient institutional rates missing in some markets. Clarify also observed that UHC initially reported uniform negotiated rates across all DRGs for some hospitals, which is likely a reporting error.

Enrichment methodologies

Clarify's all-payer claims data is used to highlight the rates that matter for payers and providers while filtering rates that will never be utilized. Clarify incorporates provider and health system characteristics and identifiers, identifying plan contracting groups, plan types, and employer affiliations, adding service line categories, estimating rates as a percentage of Medicare payments, and combining with national commercial claims samples to assess utilization. Clarify then uses a distinct approach leveraging our proprietary data sources to meaningfully assess market prices of specific services, which often number in the thousands for a single service code within a given market. Importantly, we limit rate analytics to providers with documented volume for that code. This eliminates 'zombie rates,' which are rates that are reported for providers who never have and would never provide certain services. We can also calculate volume-weighted averages and case mix adjustments across baskets of services of most interest to specific users.

FIGURE 4. CLARIFY HAS THE MOST ENRICHED RATES FIELDS IN THE INDUSTRY.



Rates enrichment

A variety of steps are taken to enrich negotiated rates, including mapping standard codes to service lines, estimating Medicare payments for those services, calculating rates as a percentage of the Medicare payment, and assessing claims volume to assess utilization and remove 'zombie rates.'

SERVICE CODE TYPES, SERVICE CLASSIFICATION, AND SERVICE LINE MAPPING

Rates are provided across a range of standard codes. These standard code types are identified by the payer and validated by Clarify using standardized logic. Standardization includes redefining or excluding codes that do not align with coding conventions, such as codes reported to be MS-DRGs that are more than three digits long. In addition to the most common codes, including MS-DRGs, HCPCS, and CPT codes, a wide range of rates associated with other code types are available in the Clarify Rates IQ Suite, as needed for specific use cases. These include APC, AP-DRG, APR-DRG, CDT, CPT, CSTM-ALL, EAPG, HCPCS, HIPPS, ICD, LOCAL, MS-DRG, NDC, PROC, RC, and UNKNOWN code types.

For rates data corresponding to an MS-DRG, CPT, HCPCS, APC, EAPG, CDT, or revenue center code, services are mapped by the standard code to one of four service classifications. These include dental, inpatient, outpatient, and unknown service classifications.

Similarly, for rates corresponding to an MS-DRG, CPT, or HCPCS, services are mapped by the standard code to one of 41 service lines. A complete list of service line mappings is available in the Appendix.

MEDICARE RATE LOGIC

Medicare rates are calculated for each standard service code in the Clarify Rates IQ Suite, primarily to calculate a percentage of Medicare benchmark for each rate. Separate methodologies are applied for inpatient, outpatient, and professional services, respectively, consistent with differences in payment by Medicare. In addition to Medicare rates estimated at the local area level described here, Clarify also calculates hospital-specific Medicare rates as described below.

Inpatient and outpatient services

Medicare payment rates for inpatient services are calculated from the Inpatient Prospective Payment System (IPPS), while rates for outpatient services are based on the outpatient prospective payment system (OPPS).

Clarify first calculates Medicare payment amounts for inpatient and outpatient codes specific to the hospital's CMS Certification Number (CCN) of the hospital. The CCN-specific inpatient rates adjust for CCN-specific wage indices and cost of living adjustments. In contrast, the DRG weights, labor share, labor rate, and non-labor rate are applied uniformly to all CCNs in a given year, consistent with Medicare payments. Similarly, calculated outpatient rates adjust for CCN-specific wage indices and the relative weights and conversion factors applied to all CCNs in a given year.

CCN provider identifiers are not supplied in publicly released rates files. To assign CCNs, Clarify maps the provided NPIs to CCNs using reference data sources. In some cases, we cannot match NPIs to CCNs necessary to calculate CCN-specific payment amounts. Because the IPPS and OPPS rates are tied to CCNs, we cannot always assign the CCN-specific calculated rates from our reference data to the enriched payer-sourced data. When NPIs cannot be assigned to a CCN, payment rates are set at the zip code level (either a five or three-digit zip code). The zip code associated with a rate is sourced from Clarify reference data based on the rate's specific NPI. These estimates aggregate all CCN-specific payment amounts in a five-digit zip code and three-digit zip code and calculate the median rate for all codes in that zip code in a year. A final Medicare rate is assigned for all relevant service codes (MS-DRG, HCPCS, or CPT) for all providers using the following hierarchy:

- CCN-specific rate for a code from most recent year available
- Five-digit zip code's median rate from most recent year
- Three-digit zip code's median rate from most recent year
- For outpatient codes only, unadjusted outpatient rates from most recent year

Examples of this hierarchy:

- An NPI-DRG combination has a CCN identified in Clarify reference data. A CCN-specific rate for the DRG is available from 2021, and a five-digit zip code median rate for the DRG is available from 2022. The CCN-specific rate from 2021 will be selected as the Medicare rate.
- An NPI-DRG combination does not have a CCN populated. A five-digit zip code median rate for the DRG is available from 2021, and a three-digit zip code median rate for the DRG is available from 2022. The five-digit zip code median rate from 2021 will be selected as the Medicare rate.

This approach is not without potential limitations, with certain Medicare payment policies not applied to allow for equitable comparison of payment rates between local providers providing the same healthcare services. Specific to inpatient payment amounts, the IPPS rate calculation used by Clarify does not account for disproportionate share hospital (DSH) adjustments or indirect medical education (IME) adjustments.

Hospitals that treat a high percentage of low-income patients may qualify for DSH adjustments, and approved teaching hospitals may qualify for IME adjustments. Regarding outpatient payment amounts, the OPPS rate is always calculated using each year's full conversion factor. Under the OPPS, CMS subjects hospitals that do not submit data on standardized quality measures to a reduced conversion factor; we do not apply this reduced factor to any provider to improve cross-provider rate comparisons. In addition, the OPPS applies payment adjustments for certain qualified hospitals (i.e., rural sole community hospitals, hold-harmless cancer centers, and pediatric hospitals) intended to promote specific CMS policy goals: calculations for this metric also exclude these adjustments. Users seeking hospital-specific Medicare rate benchmarking should utilize the adjusted Medicare rate metric which incorporates these measures.

Professional Services

Medicare rate estimates for professional services codes are based on CMS, local area-specific (i.e., carrier-locality specific) Physician Fee Schedule (PFS) pricing inputs available from CMS. The carrier-specific PFS rates are an additional source of reference data maintained by Clarify for use across the Clarify Atlas Platform. No further adjustments are made to the data, with PFS rates available from 2013 to 2023 and the most recently available year applied.
Specifically, the most recent year's non-facility PFS amount is used as a code's Medicare rate based on matching the code, carrier, and locality to Clarify's reference data. The carrier and locality associated with a rate are determined based on the rate's zip code. The PFS non-facility fee schedule amounts can vary for a code by HCPCS modifier for a year, carrier, and locality combination. Only amounts without HCPCS modifiers are selected as the Medicare rate, regardless of the modifier(s) present on a payer-sourced rate. Note that modifiers are not required to be included under the TiC rule.

PERCENT OF MEDICARE CALCULATION

Following the calculation of a Medicare payment amount for a DRG, CPT, or HCPCS code's rate, the negotiated rate amount is divided by the Medicare rate to determine the percent of Medicare. This calculation is not performed if the record is a per diem or percentage rate as indicated by its negotiated type.

Percent of Medicare rates have been used in health services research in several applications to normalize for differences in price levels and service intensity across fields.

ADJUSTED MEDICARE RATE AND ADJUSTED PERCENT OF MEDICARE

Clarify also outputs two additional benchmark metrics to use when comparing rates, "Adjusted Medicare Rate" and "Adjusted % of Medicare Rate." These are hospital-specific metrics defined as:

- Adjusted Medicare Rate - The approximate Medicare rate for a given code inclusive of hospital-specific payment adjustment (for disproportionate share hospitals (DSH), indirect medical education (IME), uncompensated care, pass through, etc.). The inclusion of hospital-specific payment adjustments is the key difference between the adjusted Medicare rate metric and the general "Medicare rate" metric described above, which is based on general Medicare payments inclusive of the wage index.

- % Adjusted Medicare Rate - The payer posted allowed amount (rate) for a given code/NPI combination divided by the adjusted medicate rate.

CLAIMS VOLUME

Despite the trillions of rows of rates data released by payers, much of this data is not associated with legitimate service volume, and so it is not meaningful for either payers or providers. Comparing rates for specific service codes by NPI with Clarify's all-payer claims data, we find substantial over-reporting of rates for provider-service code pairs with no documented utilization in claims. In some cases, these findings are obvious: dialysis clinics and clinical laboratories are not actual providers of elective surgeries despite payer-reported rates for such services. Others are less clear and show the benefit of a claims-based approach to review, such as rates negotiated with hospitals for specific services they do not actively provide.

Clarify considers multi-payer service volume, including both traditional Medicare service volume and commercial volume. Traditional Medicare service volume is estimated from 100% of Medicare claims available to Clarify, a CMS Qualified Entity; commercial claims volume is sourced from multiple proprietary commercial claims samples licensed by Clarify and applied across the Atlas platform.

MODIFIERS

Clarify has also incorporated information surrounding modifier codes directly from the payer rates files. "Modifier" is a CMS required field indicating an array of possible billing code modifiers applicable for a given rate. There are certain billing code types that allow for modifiers (e.g., the CPT coding type allows for modifiers).
If a negotiated rate for a billing code type is dependent on a modifier for the reported item or service, then payers are able to specify these modifiers via this field. We recommend only using this field on smaller, service code-specific queries and completing manual review of analyses given that modifier codes can be reported in many possible variations.

PLACE OF SERVICE

Rates are also able to be stratified with a place of service field reported directly from the payer rates files. "Place of Service" (also known as site of service) reflects the CMS-maintained two-digit code that is placed on a professional claim to indicate the setting in which a service was provided. For professional rates, this field is required per the CMS schema. This information is posted by payers as an array of possible place of service codes applicable for a given rate. Similar to modifier usage, we recommend only using this field on smaller, code-specific queries since an array can contain many possible variations.

Payer plan enrichment

ISSUER

A variety of steps are taken to enrich negotiated rates, including mapping standard codes to service lines, estimating Medicare payments for those services, calculating rates as a percentage of the Medicare payment, and assessing claims volume to assess utilization and remove 'zombie rates.'

CLEAN PAYER NAME

Clarify extracts payer name information from the rates data to assign a unique payer subsidiary for all observations. These are separated into two types of payers: employer plans and all other named payers operating under the same issuer. Additional categorization is then completed by payer, with a uniform approach implemented except where specific deviations are warranted. For example, as noted above Clarify has highlighted substantial overlap in negotiated rates within states across BCBS payers. For this reason, we have chosen to compile BCBS rates into a single database. However, we undertake steps using the payer name field to distinguish different BCBS payers' (e.g., Anthem, Regence, BCBS Texas) rates as described further below. The payer name reflects the payer who originated the rates data for a given MRF, under the assumption that this payer is likely the negotiating entity contracting these rates.

BCBS: Whose Rate is it Anyway?

Why are there Florida Blue rates in BCBS MA's MRF?

Each BCBS payer contracts with providers in its own service area but MRFs include in-network rates that extend nationwide due to the BlueCard program and other cross-payer reciprocity

Among Florida general acute care hospital rates, for example, 75% match to Florida Blue and 16% to Capital Health Plan (Tallahassee, FL) but the other 9% match to other Affiliates including Anthem, BCBS GA, BCBS MA, BCBS AL, Capital Blue, etc.

Understanding which Blue originally negotiated a rate is crucial, but identifying the contracting entity is not straightforward with the data elements presented in the MRFs alone.

Clarify has prepared a proprietary crosswalk table that maps source MRFs to their respective entity, allowing identification of rate issuers all BCBS payers.

PLAN ISSUERS

Because several BCBS plans will post the same national files, Clarify has created a plan issuers field to allow users to identify who is posting an MRF, even if they are not the entity that originated the data.

For example, if Florida Blue posted a file named 2022-12_170_22T0_in-network-rates, but this MRF has been identified as originating from BCBS Louisiana, they are a plan issuer but BCBS Louisiana is the originating entity identified in the clean payer name field. For non-national files (posted on a single site) we assign the posting entity as both the originating payer & plan issuer.

PRODUCT TYPE

Product type is derived from two data sources:

- CMS HIOS product level data available through proprietary Clarify reference data sources
- Natural-language text analysis of raw plan name values provided by hospitals and payers in their rates files

The possible output values include:

- HMO: Health Maintenance Organization plans
- PPO: Preferred Provider Organization plans
- EPO: Exclusive Provider Organization plans
- POS: Point-of-Service plans
- Indemnity: Indemnity-type plans
- Dental: Dental coverage only plans
- Other/Describe: Other plan types (including ACOs and other non-traditional plan types)
- Multiple products: Plans containing references to multiple products

An additional description of product type reference data sources is helpful to understand the latency and quality of the product type field. HIOS data can only be used to determine product type in payer-sourced rates data, as hospital-sourced rates data does not provide any HIOS information. Further, it is only possible for plans that are reported using a 10-digit (or longer) HIOS ID. Some issuers only report a 5-digit HIOS ID, and many plans (primarily employer plans) are reported using an EIN. The number of plans whose product type can be identified using HIOS data will vary from issuer to issuer.

In addition to using HIOS data, Clarify has also developed logic that searches the raw plan name value for specific keywords related to the product types, e.g., HMO, PPO, EPO, POS, and Indemnity. If, for example, a plan name matches the keyword search for HMO, the plan is assigned an HMO product type. In some cases, a plan name will contain multiple product type keywords (e.g.,"Cigna HMO/PPO"). We do not give preference to one product type over another in these instances and instead designate the plan as "Multiple products." These plans are likely a rollup of various plans with varying product types that share rates.

In payer-sourced rates data where both methods are possible for identifying product types, preference is given to the HIOS-based product type. That is, if a product type can be derived from both the HIOS data and plan for a plan, the HIOS-based product type is selected.

Limitations exist for the product type categorization as currently implemented. The fill rate for product type will vary from issuer to issuer based on a considerable variance in the raw data. Some issuers seldom report 10-digit HIOS IDs, and some use plan names that cannot convey a product type. We have identified dental plans that indicate an additional product type (e.g., a dental PPO plan). We feel that in these situations, it would be more consistent to classify the plan as a PPO product type. Note that these plans do not qualify for the "Multiple products" designation as the dental identification comes from the HIOS method, while the PPO identification comes from the plan name method.

PAYER CLASS

As a last step in our enrichment of payer data, each plan and rate are mapped to a payer class. While commercial plans are a primary interest for most users, a diverse set of payer classes are available in the Clarify Rates IQ Suite data, including Commercial, Medicare Advantage, Exchange, and Medicaid payer class types. However, Medicare Advantage and Medicaid plans are not required to be included in payer data per the TiC regulations. Most payer-sourced rates are classified as commercial or Exchange plans. We observed a wider range of payer types in provider rates, including Commercial, Dual, Exchange, Managed Medicaid, Medicaid, Medicare, Medicare Advantage, Transplant, Veterans Affairs, Vision, and Workers Compensation classes.

PLAN GROUPING METHODS

A main technical issue with payer rates data is the sheer size of payer-released files. Close to a petabyte of raw data is available for download, and this is prior to additional enrichment steps taken by Clarify. A main source of data size is the large number of raw plan names corresponding to each unique service-specific rate-NPI combination for a given payer. Many of these plan name differences do not correspond to different rates; they are "duplicates" in the sense that they do not provide any additional value or insight on rate info outside of the specific plan name. Clarify has observed examples where the same rate for the same provider map to over 40,000 unique plans offered by the same payer.

These data may be useful to some users, but consequently, fully joined data is extremely large and computationally inefficient, increasing run times and unnecessary for many users. To overcome this feature of the released data and increase computational efficiency, Clarify has prepared a schema to group plan names with the same rate into buckets to reduce data size.

This dramatically improves processing time, particularly when considered within the Clarify Rates IQ Suite. For example, if we have 40,000 rows of rate info for UnitedHealthcare at a single facility for a single MS-DRG, Clarify categorizes the plans into groups based on rate variance. We leverage product type categorizations to divide up groups by HMO, PPO, and other plan types. After reviewing rates and assigning all plans to plan groups, users can leverage two flexible approaches:

- Rates grouped at Plan Group ID level: Allowing queries at the plan group level makes insights easier to understand and faster to load.
- Rates grouped at the Raw Plan Name level: Separate tables are maintained for each payer mapping groups to plan name, as it still needs to be discoverable for some users.

Enriched rates data grouping plans result in no loss of information except for specific plan names, allowing the Suite to combine multiple payers for consolidated, cross-payer review and analysis. It allows queries to run more efficiently and provides for flat file delivery with less noise. Further drill-downs into the exact plan names may then be completed as necessary.

Provider enrichment

In addition to rate-specific and payer plan-specific enrichment efforts, enrichment of provider data found in released rates data is one additional way in which Clarify's platform standardizes and improves on publicly available data.

PROVIDER LOCATION INFORMATION

Working from released provider identifiers (NPIs) and rich, proprietary reference data, each rate is matched to a street address, city, state, county, metropolitan statistical area (MSA), zip code, locality, carrier, and census division.
This enrichment helps group data into meaningful local market comparisons across a wide variety of use cases.

FACILITY ROLL-UP, PROVIDER GROUP, AFFILIATION LOGIC

Similarly, working from released provider identifiers in the raw rates data, Clarify applies a rich mapping of provider identifiers to the organizations they operate. This includes mapping NPIs to provider names and health system characteristics.
In addition to provider group characteristics, we apply specialty information and other provider-specific characteristics to each observed rate.

TIN - ORGANIZATION NPI CROSSWALK (PAYER RATES DERIVED)

Sourcing both from proprietary reference data sources and the amassed rates data itself, Clarify can strengthen the quality of provider data by creating a TIN to organization NPI crosswalk file to fill in gaps in specific price transparency files. This allows us to allow search by both NPI and TIN across payers. While payers post rates at both the TIN / NPI levels, Clarify's provider reference and affiliation data is tied to an NPI. Therefore, provider name, specialty, state, zip, etc., will be tied to NPIs and not directly to a TIN. Therefore, it is best practice to pull in both TIN and NPI when searching by TIN, so any additional fields Clarify adds through enrichment can be viewed.

HEALTH SYSTEM

Using proprietary Clarify reference data sources, we check all providers for affiliation with a health system. Providers, including hospitals, physician groups, and individual clinicians, are assigned to either a specific health system or are given the value 'No Parent Affiliation.'

Validation and quality review

Clarify assesses data quality and regulatory compliance of payer-reported and hospital-reported rates data. We evaluate payers and hospitals on their compliance, assessing multiple metrics related to their released rates:

PAYER SOURCED

- A check to determine if the payer in question posted any MRFs per the TiC regulations
- The number of unique health systems included in the MRFs
    - Based on NPI
    - The number of unique NPIs included in the MRFs
- The number of unique MS-DRG, APR-DRG, & HCPCS codes included in the data
- With consideration given to the % of possible DRGs and a % of possible HCPCS codes
    - Do the files have all the DRGs that are listed?
    - Do the files have all the CPT/HCPCS Codes listed?
- A check to ensure the files include both inpatient, outpatient, & professional rates
- Number of unique payer names
- Recency of updates (must have at least one a year)

HOSPITAL SOURCED

- A check to determine if the hospital reported inpatient, outpatient, and professional rates
    - For Inpatient Rates: Does the hospital report negotiated rates for all required DRGs?
    - For Outpatient / Professional Rates: Does the hospital report rates for all required CPTs/HCPCS?
    - For All Rates: Does the hospital have rates associated with all the codes required under the HPTR?

Statistical analysis and trends

Clarify's analytics teams are actively working with the price transparency data to support customer-specific queries and help users better understand the data. Our team is made up of an interdisciplinary team of over 20 analysts with expertise in actuarial methods, data processing, data science, health economics, healthcare finance, health services research, and statistics. We can conduct trends analysis on our enriched rates data using service line mapping, volume data from commercial and Medicare claims, geographical attributes, and other specific features of a market, across payer plans, or for a set of providers depending on the specific use case. In general, and particularly for any national or regional analysis, there is often too much data to assess meaningful market trends unless we use our enrichment fields to inform the relevance of a specific rate. For example, lab providers reporting surgery rates without recorded experience with such procedures are completely meaningless and should be excluded from analyses. The process of identifying trends:

Combining data to the service code, provider, payer level (including the rate amount, reporting entities, etc.).

Removing the rates for codes with low materiality (typically identifying a standard market basket - using volume and rate combination).

Rolling up data at different levels of aggregation (service line, code type, NPI, etc.) for a payer using service volumes and/or market case mix of codes within service lines.

Comparing rates and rates as a percent of Medicare between payers and providers. A published example of these types of comparisons was detailed in our December 2022 research brief.

STANDARD ANALYSIS METRICS

Clarify's platform supports a range of statistical analysis methods to assess negotiated rates:

- Rate ($): The rate reported by the payer for the specified provider and code.
- Rate (%): This metric represents the contracted percentage of reimbursement for rates that are reported on a percent of charges basis.
- Rate ($, weighted by service volume): Described in greater depth on the following pages, this metric reflects the rate reported by the payer, weighted by the Combined Service Volumes. This metric is primarily applicable when viewing rates data at aggregations higher than the Standard Code level.
- Medicare Rate: As described previously in depth, this metric represents the Medicare rate based on the various factors that determine the rate (e.g., site of care, CCN, zip code).
- % of Medicare Rate: The payer posted negotiated rate for a given code/NPI combination divided by the Medicate rate.
- Adjusted Medicare Rate: The approximate Medicare rate for a given code inclusive of hospital-specific payment adjustment (for disproportionate share hospitals (DSH), indirect medical education (IME), uncompensated care, pass through, etc.).
- % of Adjusted Medicare Rate: The payer posted allowed amount (rate) for a given code/NPI combination divided by the adjusted Medicate rate.
- # of Distinct Prices: The unique number of rates published between a payer and provider for a specific code. Reasons for multiple distinct prices may be related to different negotiated rates for various product types (e.g. HMO, PPO, and EPO) or if the provider is included in multiple networks.
- # of Prices: The number of rates reported by a payer for a specific provider and service code.
- Medicare Service Volume: The number of Medicare-reimbursed services rendered by the selected provider for the selected basket of codes. This information is sourced from QE claims, which represent 100% of Fee-for-Service Medicare services for all beneficiaries. Actuarial techniques are applied to the data to assess claims maturity to ensure that volumes are not understated due to claims incurred but not yet reported (IBNR).
- Commercial Service Volume: The number of reimbursed services rendered by the selected providers and codes. This information is sourced from various proprietary data sources, and utilizes actuarial techniques to assess claims maturity and ensure that volumes are not understated due to IBNR.
- Combined Service Volume: The combined amount of the Medicare Service Volume and the Commercial Service Volume metrics.
- Unit Rate Benchmarks: Specific to the Provider Sourced Rate Intelligence module, rates data can be benchmarked by various dimensions (e.g., service line, code, hospital, payer) to assess median, mean, upper/lower quartile, minimum and maximum rates as a percent of Medicare.

VOLUME-WEIGHTED RATE METRICS

As described previously, Clarify incorporates national all-payer claims data to match negotiated rates to service code volume at the provider–service code level.

Using this volume data, Clarify has developed a volume-weighted rate metric. This metric is crafted to incorporate claims volume into rate aggregations, providing a more comprehensive and valuable view by combining price transparency data with the extensive information available in Clarify's all-payer claims. It aims to ensure a more comprehensive understanding by factoring in utilization while comparing rates at different levels in the market or across service lines.

The core of this metric revolves around the combined commercial and Medicare volume associated with a specific code/NPI combination. This volume is utilized to weight a rate as it gets aggregated to broader levels, such as service line, hospital, or the entire health system. To ensure accurate and unbiased aggregations, the metric also considers the number of distinct prices linked to the given NPI/code. This safeguards against potential discrepancies arising from duplicated rates in the raw files.

The calculation of the volume-weighted rate is methodical. First, for a particular code/NPI combination, the rate is multiplied by the volume. This resultant value is then divided by the count of unique rates associated with the same NPI / standard code combination. Taking this quotient, you set it as the numerator and divide it by the combined volume. Lastly, to obtain the final volume-weighted rate, you divide the result once more by the count of distinct rates. Mathematically, this can be represented as:

$$\Sigma(\text{rate} * \text{combined service volume for associated standard code/count of distinct rate for that standard code})$$

$$\Sigma(\text{combined service volume for all standard codes} / \text{combined service volume for all standard codes})$$

Users have the flexibility to incorporate this metric into their reports using the Payer Rates UI, especially when analyzing rates at broader categories like service line, hospital, or health system levels. It provides an aggregated rate that includes utilization for every service and NPI that's part of the aggregation. However, if users are inclined towards metrics not adjusted by Clarify's utilization data, they can still access and use the traditional rate metrics. These metrics offer straight averages of distinct prices as found in machine-readable files.

It's important for users to understand the current limitations of the "Volume Weighted Rate." These metrics aren't compatible with "Percentage" or "Per diem" negotiation types, given that these values don't aggregate with conventional dollar amounts typically negotiated on a contract. Similarly, any calculations involving a percentage of Medicare metrics will default to the non-weighted rate metric. Separate volume-weighted metrics tailored for these contract types may be discussed with our analytics team.

FILTERING OPTIONS

Clarify's platform also provides a range of filtering options to narrow review to the rates that matter most for users and allowing focus on rates with full coverage, including by filtering to:

- % of Medicare Rate > 0
- Medicare Rate > $0
- % of Medicare metric falls between 150-700%
- All Breakouts Include a Rate
- Combined Service Volume > 0
- Commercial Service Volume for Provider > 0
- Medicare Service Volume for Provider > 0
- Rate Amount > $0
- Rate Percentage > 0%
- Filtering only to specific TINs, NPIs, or states of interest

FUTURE CONSIDERATIONS

Price transparency data and Clarify's approach to payer- and hospital-sourced data are continue to evolve and improve over time. The Clarify team is currently refining approaches to reflect:

- Payer plan network identification leveraging proprietary network reference data sources
- Opportunity analyses for payers and providers allowing consideration of hypothetical changes in negotiated prices
- Integration of payer and provider rates data for cross-validation, gap filling, and identification of Clarify 'best' negotiated rate within a local market for a given service
- Time-series analysis of rates changes over time

Appendix

DATA DICTIONARY

The current version of our data dictionary is available upon request.

SAMPLE SQL QUERIES

Some example SQL queries for use in starting out with our rates flat files are available upon request.

COMPARISON OF HOSPITAL AND PAYER PRICE TRANSPARENCY REGULATIONS

	HOSPITAL PRICE TRANSPARENCY FINAL RULE	TRANSPARENCY IN COVERAGE FINAL RULE
EFFECTIVE DATE	JANUARY 1, 2021	JULY 1, 2022
DATA REQUIREMENTS	Machine-readable files (MRFs) should include all individual items and services (including service packages) provided by hospital: The gross charge (the charge for an individual item or service that is reflected on a hospital's chargemaster, absent any discounts). The discounted cash price (the charge that applies to an individual who pays cash, or cash equivalent, for a hospital item or service). The payer-specific negotiated charge (the charge that a hospital has negotiated with a third-party payer for an item or service). The de-identified minimum negotiated charge (the lowest charge that a hospital has negotiated with all third-party payers for an item or service). The de-identified maximum negotiated charge (the highest charge that a hospital has negotiated with all third-party payers for an item or service).	MRFs containing the following sets of costs for items and services: In-Network Rate File: In-network machine-readable files include negotiated rates for all covered items and services at the plan level, for all medical codes, for contracted rates between the plan or issuer and in-network providers. Out-of-network Rate File: Out-of-network machine-readable files include allowed amounts for covered items, services based on billed charges and allowed amounts including historical amounts. Files are not required if there are fewer than 20 claims for a service for a provider. Prescription Drug File: CMS delayed enforcement and then rescinded that delay. No data schema updates have been provided.

	HOSPITAL PRICE TRANSPARENCY FINAL RULE	TRANSPARENCY IN COVERAGE FINAL RULE
REFRESH SCHEDULE	YEARLY	MONTHLY
FORMAT	Standard schema requirements by CMS beginning in 2024.	Must follow a defined schema to allow for data collection, consolidation, and parsing.
PENALTIES & ENFORCEMENT	$109,500 - $2,007,500 for a full calendar year of noncompliance per hospital. Only two fines issued to date, some warning letters also issued.	$100 per day per each affected member. No active enforcement to date.
COMPLIANCE	Estimated between 25 – 70 percent in full compliance prior to 1/1/24 requirements.	All national and regional payers posting data, with some payers reporting incompletely.
EFFECTIVE DATE	January 1, 2021 (with more updates in 2024 and 2025)	July 1, 2022

Clarify's rate service line categories

Administered Drugs
Ambulance
Ancillary Services
Anesthesia
Behavioral Health
Cardiac Services
Durable Medical Equipment
Emergency Department
Endocrinology
Enteral and Parenteral Therapy
Evaluation Consultation
Gastroenterology
Head Neck
Hematology/Oncology
Home Health Services
Infectious Disease
Lab
Medical and Surgical Supplies
Nephrology
Neuroscience
Obstetrics
Ophthalmology
Orthopedics
Orthotics and Prosthetics
Other
Other Diagnostic Study
Other Therapeutic Procedures
Pathology
Performance Measurement
Physical Therapy
Pulmonology
Rehab Aftercare
Skin and Wound Care
Temporary Codes
Trauma
Unknown
Urology
Vaccination
Vision Hearing and Speech Supplies
Women's Health

Glossary

ATLAS PLATFORM
Clarify's analytics platform is designed to provide healthcare price transparency and enriched claims data, insights, and tools to help stakeholders, including both payers and providers make data-driven decisions and better understand healthcare spending and negotiated rates.

BUNDLED PAYMENT
A single, comprehensive payment made to healthcare providers for a group of related services, such as those associated with a specific medical condition or procedure, to encourage coordinated care and cost containment. Available in price transparency data for select healthcare services.

FEE-FOR-SERVICE PAYMENT
A traditional healthcare payment model in which providers are reimbursed for each individual service or procedure they perform, rather than receiving a fixed or bundled payment for a patient's overall care. The best, readily available source of pricing information in price transparency data comes from fee-for-service negotiated rates.

HOSPITAL PRICE TRANSPARENCY RULE
A regulation implemented by the Centers for Medicare & Medicaid Services (CMS) requiring hospitals to publish their standard charges for items and services, including negotiated rates with payers, in a machine-readable format.

IN-NETWORK PROVIDER
A healthcare provider who has a contractual agreement with a health insurance plan to provide services at a negotiated rate.

INSTITUTIONAL SERVICES
Healthcare services provided by facilities such as hospitals, nursing homes, or ambulatory surgical centers, encompassing a wide range of treatments, diagnostics, and therapeutic interventions.

NEGOTIATED RATE
The price agreed upon between a healthcare provider and an insurer for specific services, usually lower than the standard charge.

NPI (NATIONAL PROVIDER IDENTIFIER)
A unique 10-digit identification number issued by the CMS to healthcare providers for use in administrative and financial transactions, such as billing and claims submission, which primarily identifies payers in price transparency data.

PERCENT OF CHARGES PAYMENT
A payment methodology in which reimbursement for healthcare services is based on a percent of the provider's standard billed charges. This approach can result in varying payments depending on the provider's charge structure and negotiated rates with insurers.

PER DIEM PAYMENT
A fixed daily rate paid to healthcare providers, such as hospitals or long-term care facilities, for each day a patient receives care, regardless of the specific services provided.

PROFESSIONAL SERVICES
Healthcare services provided by licensed practitioners, such as physicians, nurse practitioners, or physician assistants, in the diagnosis, treatment, and management of patients.

TIN (TAXPAYER IDENTIFICATION NUMBER)
A unique identification number assigned by the Internal Revenue Service for tax reporting purposes, which can be used to identify healthcare organizations and providers in financial transactions and are reported instead of NPIs in the price transparency data by some payers.

Clarify

About Clarify Health

Clarify Health Solutions® is a healthcare data and analytics company trusted by some of the most established organizations in healthcare, including providers, payers, tech and services, and life sciences. The Clarify Atlas Platform® is the foundation, leveraging the industry's largest and most robust dataset to map over 300M+ lives to deliver 20B+ AI-powered predictions to surface actionable insights with unparalleled speed and precision.

To learn more about how Clarify's data and analytics solutions can help your organization answer critical business questions about provider performance, visit clarifyhealth.com.