simple healthcare

VIA ELECTRONIC SUBMISSION

February 23, 2026

Centers for Medicare & Medicaid Services
Department of Health and Human Services
Attention: CMS-9882-P
P.O. Box 8016,
Baltimore, MD 21244-8016

Internal Revenue Service
Department of the Treasury

Employee Benefits Security Administration
Department of Labor

Re: Transparency in Coverage (File Code: CMS-9882-P)

Dear Administrator Oz, Acting Commissioner Bessent, and Assistant Secretary Aronowitz:

Simple Healthcare is committed to price transparency in US healthcare and has extensive experience with Transparency in Coverage (TiC) policy and data. We have provided policy recommendations to improve TiC data including actionable recommendations to make the data more usable and highlighting the need for data quality standardization.

We have extensive experience analyzing TiC data and have used it to answer a variety of policy-relevant topics such as quantifying the prevalence of 'ghost rates' or negotiated rates for billing codes corresponding to provider groups who are unlikely to bill for that code, assess the data quality by investigating the level of data completeness for the major national payers and demonstrate how provider market shares are associated with higher negotiated rates.

We are grateful for the opportunity to submit comments on the proposed Transparency in Coverage rule and are in broad support of the Departments' proposed direction to improve TiC. Our primary recommendations focus on making TiC data easier to use for all stakeholders. They are:

1. Including Procedure Volume: **We advocate for including procedure volume at the provider group level** for every network. Information on procedure volume can inform quality, help in assessments of network adequacy, and help improve data quality.

2. Removing "ghost rates": **We recommend using a combination of procedure volume at the provider-group level and specialty-based taxonomy to remove ghost rates** and improve TiC

data quality. Specialty-based taxonomy on its own is insufficient and will lead to exclusion and inclusion errors.

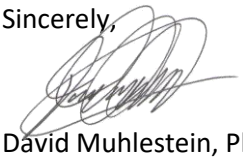
3. Bundled Prices: **We recommend that specific requirements be built around bundled and pass-through payments** to ensure that full pricing information is available.

4. Relational CSV (comma-separated values) file format: **Insurers should be required to release their machine-readable files in a relational, rectangular format.** This will result in a manageable file size as the data is spread across multiple tables or files. It will make the data easier to access and use in comparison to JSON (JavaScript Object Notation) files.

We provide more information and specific recommendations below. We also are attaching a paper that goes into some of these topics, including more specific detail on the implications for ghost rates, how to assess bundles and the implications of working with JSON vs a relational, rectangular data format.

The proposed rule is a large step in the right direction of increasing price transparency. We appreciate the opportunity to comment on this proposed rule and welcome the opportunity to assist the Department on topics we discuss. For questions, please contact me directly at david@simplehc.com.

Sincerely,



David Muhlestein, PhD JD
CEO, Simple Healthcare

Introduction

Simple Healthcare is a company that is heavily involved with price transparency, including working with current transparency data to evaluate the broader health system and help individual companies operate more effectively. Simple Healthcare has extensive experience working with Transparency in Coverage (TiC) files, and we have led several studies using TiC data with relevance for policy, academia, and users of TiC data writ large.

We are submitting this comment letter in response to the proposed TiC rule by the Internal Revenue Service, the Employee Benefits Security Administration, and the Department of Health and Human Services (Departments) on December 23, 2025.

Simple Healthcare strongly supports the proposals in the rule that focus on improving the TiC data quality. Emphasis on standardizing the machine-readable files and improving the accuracy and accessibility of the data will go a long way in making TiC data easier to use for interested parties in policy making, industry, academia and others.

We strongly support moving to address excessive data that has negotiated rates for providers who do not perform those services (ghost rates) but have recommendations to improve upon the proposed rule. We also strongly recommend that volume data be shared at the group level within the negotiated rate files, while adopting a more nuanced utilization file that only includes binary participation at the individual provider level.

Background and Experience with TiC files

Simple Healthcare has worked extensively with the TiC machine readable files since their initial release in 2022. This work has included parsing, normalizing, and processing thousands of files end-to-end, as well as systematically evaluating the completeness, usability, and integrity of the files across issuers and markets. Through this process, Simple Healthcare has identified recurring issues with TiC files, such as inconsistencies in file structure, incomplete or missing information, presence of ghost rates, and other issues that limit the quality and usability of the data. These findings have been documented in various peer-reviewed and industry publications:

- *High Prevalence of Ghost Rates in Transparency in Coverage Data* reveals that a vast majority of negotiated rates in TiC files are “ghost rates”, or rates for services that a provider is unlikely to ever perform. Specifically, across major payers, about 91.8% of all reported negotiated rates fell into this category, undermining the usability of the data without substantial cleaning. (https://academic.oup.com/healthaffairsscholar/article/3/11/qxaf212/8321476)

- *Price Transparency With Gaps: Assessing the Completeness of Payer Transparency in Coverage Data*. This analysis of 2025 TiC files from major national insurers showed that across all payers, a substantial amount of rates was missing, or in other words, TiC data was incomplete. The physician and hospital outpatient price data was generally more complete than the inpatient data. (https://www.ajmc.com/view/price-transparency-with-gaps-assessing-the-completeness-of-payer-transparency-in-coverage-data).
- *Improving Price Transparency Data: Recommendations From Practice*. This paper highlights key practical challenges in working with raw price transparency data, such as inconsistent formatting, large file sizes, and missing elements. It offers actionable recommendations including better standardization and quality control to make the data more usable for employers, researchers, and policymakers. (https://www.healthaffairs.org/content/forefront/improving-price-transparency-data-recommendations-practice).
- *Hospital System Market Share and Commercial Prices: a Cross-sectional Approach Using Price Transparency Data*. This paper uses price transparency data from three large, national insurers, and finds that a one percentage point increase in hospital system market share is associated with an $88 to $118 higher negotiated rate per admission. All else equal, a hospital that is part of a system with a 10-percentage point higher market share can expect from $880 to $1,180 more per admission relative to a hospital with lower system market share (5.4% to 6.2% of the median price). (https://link.springer.com/article/10.1186/s13561-024-00580-w).
- *Commercial Insurer Price Transparency: A Comparison Of Four National Payers*. This study finds significant disparities in how four major commercial insurers disclose negotiated rates, with substantial variation in completeness and format that limits meaningful cross-payer price comparisons. Findings suggest that standardization in data reporting and improved data quality are needed to increase the use of TiC data. (https://www.healthaffairs.org/content/forefront/commercial-insurance-price-transparency-comparison-four-national-payers).
- *Improving Transparency in Coverage Data: Reducing Ghost Rates, Adding Utilization, and Standardizing File Structure*. This white paper provides technical recommendations to help strengthen the proposed Transparency in Coverage rule from the Centers for Medicare & Medicaid Services by improving how price transparency data is structured, reported, and filtered. Drawing on empirical analysis of current TiC data, the paper outlines 17 practical, implementation-ready recommendations across ghost rate filtering, utilization reporting, bundled payment disclosure, and machine-readable file architecture. The goal is to help ensure transparency data is accurate, interpretable, and usable for real-world decision-making by employers, researchers, and policymakers, while remaining feasible for payer implementation using data already maintained within existing systems. (https://cdn.prod.website-files.com/687bd1da8a52d05033fdf6ed/6997e7813f3838b5850311a8_Improving%20TiC%20Data%20White%20Paper.pdf).

Priority issues and recommended focus areas

TiC data has the potential to be highly valuable for employers, researchers, consumers, and plan fiduciaries. Employers can use information on negotiated rates across plans and networks to fulfill their fiduciary obligations to employees by selecting the most cost-effective coverage options. In addition to reducing overall healthcare spending for the employer, this could potentially lead to lower premiums and out-of-pocket costs for employees, thereby directly benefiting consumers.

Researchers and policy analysts can use the data contained in TiC files to deepen our understanding of healthcare markets. The data make it possible to identify price variation across markets, services, and providers, examine patterns of higher prices, and analyze the structural and competitive factors that drive those differences.

However, to fully realize these benefits, TiC files must be made significantly easier to access, interpret, and use. In particular, additional and improved data fields are needed, including:

- **Consistent and transparent specialty-based taxonomy that solves the issue of ghost rates without being over- or under-inclusive by including or excluding providers who are unlikely to provide a service, respectively.**
- **Meaningful utilization and volume data to contextualize negotiated rates and assess market significance.**
- **Complete billing information for bundled services or procedures**
- **A clear and accurate definition of what constitutes a network to ensure comparability across insurers.**
- **Consistent identifiers for networks and standardized units for negotiated rate type, for example.**

Excluding unlikely provider-to-service mappings; ghost rate reduction

We support the Departments' objective to reduce file size and improve usability by removing implausible provider-to-service rate combinations. However, we caution that specialty- or taxonomy-based exclusion alone will not reliably achieve this goal. Evidence from current TiC data demonstrates that taxonomy-only approaches can be simultaneously over-inclusion and under-inclusion because specialty labels are imperfect proxies for actual service delivery patterns. Claims-based utilization

patterns show that many providers within a specialty do not perform a given service, while adjacent provider types outside the expected specialty frequently do.

In the near term, a plan-specific inclusion approach anchored in claims experience is likely more realistic than a national taxonomy mapping. Payers already maintain claims data and can implement group-level inclusion rules using existing infrastructure. Over the long term, we support the Centers for Medicare & Medicaid Services (CMS) developing national service-to-specialty mapping frameworks to improve consistency across issuers, but these mappings should be informed by observed utilization patterns rather than static taxonomy assumptions.

In a new publication we released, using code 27447 (knee replacement), we demonstrate how specialty/taxonomy exclusion can be both over-inclusion and under-inclusion.

For code 27447, orthopedic surgeons bill for 60.8% of claims, while physician assistants bill for 33.3%, and nurse practitioners bill for 4.6% (all other specialties represent 1.3%). Of all orthopedic surgeons, only 32.1% bill for 27447, while only 4.3% of physician assistants and 0.3% of nurse practitioners bill for this. The implication is that no matter which taxonomy you choose, there will either be over-inclusion, or under-inclusion.

Over-Inclusion: The most likely taxonomy to include for 27447 is orthopedic surgeons, but only 32.1% perform this, so including all orthopedic surgeons retains large numbers of codes for orthopedic surgeons who do not perform knee replacements.

Under-Inclusion: Excluding non-physician providers removes a substantial portion of real utilization. For example, physician assistants account for a meaningful share of knee replacement claims and would be excluded under physician-only logic.

In Table 1, Simple Healthcare's analysis shows how specialty-based taxonomy can lead to both over-inclusion and under-inclusion for billing code 27447 (Total Knee Arthroplasty). For example, for UnitedHealthcare, if you only include billing codes for providers that actually have performed 27447, then only 4,101 of the reported 46,352 rates would be included. However, if you included rates for all orthopedic surgeons, 5,993 rates would be included (meaning that approximately 32% would be ghost rates). If you include Physician Assistants (PAs) and Nurse Practitioners (NPs) with the inclusion criteria, then there would be 19,246 rates, meaning that 79% would be ghost rates.

Table 1: Count of Negotiated Rates for Billing Code 27447 by Insurer and Inclusion Criteria

Insurer	Reported Rates	Rates Based on NPIs in Claims (volume)	Rates Including an Orthopedic Surgeon	Rates Including an Orthopedic Surgeon or PA/NP
Aetna	31,368	4,982	5,638	10,067
Cigna	15,399	2,073	1,895	5,446
UnitedHealthcare	46,352	4,101	5,993	19,246

Similar things happen with other billing codes, particularly those that are performed relatively infrequently by a variety of different specialties. Take code 83540 (Iron Serum Test), for example. This is billed by dozens of specialties. The most common – hematology-oncology – bills 29.8% of the claims for this, but only 11.4% of hematology-oncologists ever bill for it.

Table 2: Count of Negotiated Rates for Billing Code 83540 by Insurer and Inclusion Criteria

Insurer	Reported Rates	Rates Based on NPIs in Claims (volume)	Rates Including One of the most common Specialties
Aetna	23,338	2,563	22,224
Cigna	2,722	112	2,141
UnitedHealthcare	43,450	2,073	36,855

Looking at UnitedHealthcare, there were 43,450 reported negotiated rates for billing code 83540. If you limit the rates based on providers who have previously billed for that code, there would be 2,073 rates.

If you include all of the specialties that commonly bill for that code, there would be 36,855 rates, meaning that 94% of the rates would be ghost rates.

A specialty-based rule – a single specialty-based rule or a multiple specialty rule – cannot by itself achieve the objective of cleaning the TiC data of ghost rates in a meaningful way.

We recommend that CMS require rate inclusion to be based on a combination of claims-based volume and specialty-based taxonomy. Anchoring inclusion to real utilization reduces the chances of both over- and under-inclusion, improves reproducibility across issuers, and relies on data already available within existing payer systems. This approach is operationally feasible as insurers already maintain claims databases for their internal purposes.

Volume should be reported at the provider group level and can be presented in three categories: 0 claims, 10 or fewer claims, and 11 or more claims. This would be consistent with the existing federal public reporting patterns that suppress very small cells but do not discourage meaningful analyses.

For provider groups reporting 0 claims, we recommend using specialty-based inclusion criteria if they are highly likely to provide the given service. This forward-looking inclusion is important to avoid not including data on groups that are newly contracted with the insurer, newly trained providers, or providers expanding into new service lines. For example, a newly trained orthopedic surgeon may not have performed a knee replacement during the lookback period but is highly likely to perform one during the next reporting period. Allowing inclusion for high-probability specialties preserves legitimate network representation while still removing most ghost rates. This does lead to some over-inclusion, but we feel this small over-inclusion is better than systematic under-inclusion of these other groups. The specialty-based criteria might include specialties where at least 20% bill for the procedure, but the actual number could be determined through a formal analysis.

We recommend that CMS specify a standardized timeframe and level of inclusion to ensure consistency across insurers. Specifically, we recommend that inclusion should be based on whether the provider group billed for the service within the previous twelve months of claims data. **Inclusion should be evaluated at the provider group level (Tax Identification Number (TIN)-level) rather than at the individual National Provider Identifier (NPI) level**. Group-level inclusion better reflects how services are delivered and contracted in commercial networks and avoids the multi-counting and attribution challenges that occur when utilization is measured at the individual provider level.

Utilization File and procedure volume reporting

Information on procedure volume is essential to inform policy and conduct rigorous analyses. In the absence of alternative measures for quality, volume of procedures conducted by providers remains the best available proxy for quality. Volume data is also critical to support and assess network adequacy assessments. For example, without data on procedure volume, it would take guesswork to assess network adequacy for knee replacements using the TiC data if an orthopedic group included in the data only provides services for hand and wrist procedures. Inclusion of provider group and procedure level information on volume will significantly enhance the quality of TiC data by allowing for a relatively straightforward identification of 'ghost rates' or rates for services that are unlikely to be provided by the corresponding provider group.

We recommend that volume should be reported for every network and it should be organized by provider groups. A provider group is defined as a combination of one or more TINs, each associated with one or more NPIs. These groups may have different negotiated rates that vary based on known factors such as place of service, billing code modifiers, or provider specialty within the plan. Procedure volume should be attributed to a provider group as long as one or more provider NPI within the provider group bills for that procedure, or the group bills for that procedure.

NPI information in the claims data is not straightforward to use and can be confusing as the performing NPI, billing NPI, and referring NPI can be different providers. For example, an entire provider group may bill under a single NPI, or multiple NPIs may appear on claims for the same services, leading to potential double-counting. Counting procedure volume by NPI is also not comparable across different payers or different claims data systems.

We recommend that to address these challenges, the proposed "Utilization File" rules should be adopted. NPI-level utilization should be retained as a binary indicator, defined as having at least one reimbursed claim as the **performing NPI**, to support the identification of individual providers who perform specific services. This utilization indicator should be reported at the network level, by NPI, and should include information on billing code type, billing code, place of service, billing code modifiers, and other major fields typically found in claims datasets.

The reporting structure for procedure volume should be aligned with CMS claims release policies. Provider groups with no reported cases may be listed as having zero volume, groups with between zero and ten cases should be categorized as "10 or fewer," while groups with eleven or more cases should have the actual number of cases reported.

We recommend that volume data should be included in quarterly releases, and the volume should reflect the time period from approximately 18 months to six months prior to the data release. For example, if a file is released on January 1, 2028, the reported volume should cover the period from July 1, 2026, through June 30, 2027.

Network reporting, identification, and naming

Effective implementation of the TiC rule requires a clear and consistent definition of what constitutes a distinct provider network. **Our recommendation is that networks should be defined as a distinct combination of providers and negotiated rates. Using brand or marketing names can obscure meaningful differences, limiting the conclusions that can be drawn from the data**. In practice, multiple insurance products may use the same underlying network but differ in benefit design elements, such as cost-sharing or out-of-network coverage, and network naming conventions vary widely across insurers (e.g., market-facing names versus internal identifiers). As a result, plan marketing labels alone, such as similarly named "Choice" or "Choice Plus" products, are insufficient for distinguishing networks in the data.

To improve clarity and usability, CMS should require stable, standardized network identifiers, along with explicit linkage between insurance products and the network ID that they use. Further, the Table of Contents (ToC) file should include links to standardized plan-level information about the plans that connect to each network, enabling users to understand when multiple plans with different benefit design elements reference same underlying network structure. For example, Aetna has multiple plans with various benefit design elements that all reference the same network: HSA Aetna Choice POS II, Aetna Choice POS II, Open Access Aetna Select, etc. Thus, the ToC file will have the marketing network name. **Examples of standard identifiers include the Health Insurance Oversight System (HIOS) number and the group number that shows up on insurance cards**, that would help consumers (and the companies that help them) match their specific insurance to their network.

Bundled and Pass-through Payments

We recommend including information that can be leveraged to calculate the cost of bundles of services that might be billed with or without adjustments to the base rates. In its current form, TiC data does not include data for such bundles, and neither does it contain sufficient information that can enable estimation of the cost of bundles.

In our recent publication, we explore how bundled payments, particularly those performed at Ambulatory Surgery Centers (ASCs), remain inconsistent across files and contracts by using Total Knee

Arthroplasty as an example. Table 3 and Table 4 below help illustrate this point. Table 3 shows an example of what the total cost for a Total Knee Arthroplasty might be. But estimating the total cost of this bundle from the TiC data is almost impossible as the data available is not sufficient, as shown in Table 4.

Table 3: An Example of a Bundle of Services for Total Knee Arthroplasty

Billing Code	Billing Code Modifier	Description	Category	Price
27447		Total knee arthroplasty	ASC Facility Fee	$18,500
27447		Total knee arthroplasty	Surgeon Professional	$3,200
27447	80	Total knee arthroplasty - assistant surgeon	Assistant Surgeon	$650
01402		Anesthesia for total knee replacement	Anesthesia	$1,100
C1776		Joint device (implantable)	Implant/Device	$5,500
76942		Ultrasonic guidance for needle placement	Imaging	$150
64447		Femoral nerve block	Anesthesia Add-on	$425
J2001		Lidocaine injection	Medication	$18
J1100		Dexamethasone injection	Medication	$32
J2405		Ondansetron injection	Medication	$25
			Total	$29,600

Table 4: An Example of what Current TiC Data Contains for a Bundle of Services for Total Knee Arthroplasty

Billing Code	Billing Code Modifier	Description	Category	Price
27447		Total knee arthroplasty	Surgeon Professional	$3,200
01402		Anesthesia for total knee replacement	Anesthesia	$1,100
76942		Ultrasonic guidance for needle placement	Imaging	$150
64447		Femoral nerve block	Anesthesia Add-on	$425
			Total	$4,875

Insurers should be required to publish their bundle definitions and methodologies. When grouping tools or "groupers" are used to construct bundles, health plans should also be required to disclose the specific groupers they rely on. In addition, regulators should specify a minimum set of metadata fields that must be included for each bundle, such as a bundle identifier, bundle name, underlying logic, grouping methodology, component billing codes, and site-of-service applicability.

In areas where Medicare already defines standardized rollups, such as Ambulatory Payment Classifications (APCs) for outpatient services, existing established methodologies can serve as reference points. However, commercial bundling practices vary widely, particularly for ambulatory surgical centers and many surgical procedures, and not all commercial plans rely on APCs or Diagnosis-Related Group (DRGs). This variation makes it difficult to compare bundled prices across issuers.

To improve transparency, plans should be required to release full information on expenses that may not be associated with standard billing codes. This includes costs such as medical devices and other items that are reimbursed through pass-through payments or separate cost-based arrangements.

Data elements and integrity issues to address

Percent of billed charges

Negotiated rate when provided as percent-of-billed-charges is not meaningful since the exact billed amount is not disclosed. The value of negotiated rates in the form of percentages cannot be compared to the negotiated rates that are provided as dollar amounts.

We recommend that insurers be required to provide the average paid amount along with the percentage number. The billed amount should be the average billed amount specific to the provider group.

Non-reimbursable codes

To improve interpretability of the TiC data, non-reimbursable billing codes should be consistently flagged. Some issuers may report codes that are used solely for administrative or eligibility verification purposes and do not represent a meaningful reimbursable price, often appearing as nominal amounts (e.g., $0.01).

Without explicit indicators, these entries can be misinterpreted as valid negotiated rates. The Departments should therefore require dedicated fields that identify whether a code is non-reimbursable and specify the reason for its inclusion, such as informational-only use, internal processing, or service packaging, so that users can accurately distinguish between actionable price information and administrative artifacts in the data.

Units of analysis / unit definitions

To ensure that all claims-based amounts are interpretable and comparable, all reported amounts should include an explicit unit of measure or unit basis. Without standardized units, such as per service, per day, per stay, per month, or per episode, even clearly reported rates can be misleading or unusable. For example, certain billing codes (e.g., K0606) may be billed on a monthly or annual basis, and in the absence of any information about corresponding unit, the reported amount cannot be meaningfully interpreted. This issue will become increasingly critical as CMS implements forthcoming prescription drug transparency requirements, where unit definitions are essential to understanding price variation. **We therefore recommend that CMS require a unit of measure to be reported alongside each billing code and define a standardized set of allowable units to promote consistency, usability, and downstream analysis across payers and markets.**

Common names

The requirement for insurers to add the common provider network name to the In-network rate file needs to be more specific. An insurer could have a common name for internal purposes that is different from what is commonly called in the market. We recommend that insurers be required to add two fields – the common name and an identifier associated with the network that is unique to the said common name. Insurers should also specify if the common name is the one that is commonly used in the market or internally by the insurer to identify the network.

Provider information

Currently, TiC data contains information on organization TINs and provider NPIs, but it is subject to inaccuracies when matching providers to geographical locations. For example, a TIN based out of one state will have provider NPIs from a different state. **This can be resolved by requiring plans to release an NPI level data set containing provider name, specialty and address on an annual basis. This could be incorporated into the taxonomy file.**

Average paid amounts

In many cases, negotiated rates alone do not fully reflect the amounts typically paid due to contractual adjustments, bundling, or other payment mechanisms that are not readily interpretable within standard TiC schemas. Requiring plans to report average paid amounts would provide a more accurate and accessible measure of typical charges, enabling analysts, employers, and individuals to better understand typical charges**. We recommend that CMS require plans to include average paid amounts for each billing code and provider group for which the insurer has made payments to in the prior year.** This approach would improve the interpretability of payer data and align the TiC requirements with the revised hospital price transparency standards, which already mandate the reporting of average amounts. **When payments are made as a bundle where a single payment includes multiple underlying billing codes, then a derived price can be used for code-level estimates, following a similar methodology already used to calculate derived prices.**

Making the data more accessible and discoverable

A standardized approach should be adopted to make machine-readable price transparency files easier to discover and access.

To make the data easy to discover and access, **we recommend a requirement for health plans and insurers to submit the location of their TiC file postings, including the root directory URL, to CMS.** This

would enable CMS to maintain and host a centralized directory containing the location of each plan's data. This centralized directory would only provide links to where the data can be found and not the data itself.

To ensure ease of discovery and continued access, plans should be required to update this information on a regular basis, or as and when there is a change in the location of their TiC data. **We strongly support the Departments' proposal to require a standardized text file on each plan's homepage** containing this information. In addition, plans should be required to include a clearly visible link to their price transparency information in the website footer similar to existing requirements for the hospital price transparency data.

Compliance and enforcement observations

While machine-readable files may be technically posted in compliance with TiC rule requirements, they can still lack the completeness or data integrity necessary for real-world use.

To address this gap, we recommend that CMS should adopt a tiered compliance framework that distinguishes between basic posting and discoverability, schema validity, data completeness, and data integrity. Establishing clear standards for assessing whether plan data are complete, rather than merely present, would enable more consistent oversight and provide users with data that is more actionable and usable.

Evidence from prior analyses of TiC data completeness demonstrates that many files meet technical requirements while omitting critical information needed for meaningful interpretation and comparison. For example, in a 2025 analysis of the completeness of TiC data (*Price Transparency With Gaps: Assessing the Completeness of Payer Transparency in Coverage Data*), we found that across major payers, negotiated rate completeness is often highest for physician specialties and lowest for inpatient files. Completeness varied across payers, as UnitedHealthcare's physician groups were near complete, but inpatient data were sparse. Cigna showed high completeness for physician specialties and inpatient data, but limited hospital outpatient rates. Aetna demonstrated moderate to good physician completeness, midrange outpatient hospital data, and heterogeneous inpatient results. Overall, physician and hospital outpatient data were typically usable; inpatient data were insufficient. These findings suggest that more needs to be done to ensure that TiC data is complete and usable.

Non-compliance examples

Currently, many examples of non-compliance exist in TiC data, particularly with respect to provider identification and grouping. In practice, some issuers have published files that lack usable TIN-level or provider group-level information. For example, there are cases where a single issuer-reported TIN is

used across all reported provider groups, eliminating the ability to distinguish negotiated rates at the group level (example: SelectHealth), as well as instances where provider TINs are masked entirely (example: University of Utah). Additionally, issuers vary widely in how they associate NPIs with TINs and provider groups, resulting in inconsistent provider group structure or definition across files and limiting comparability across markets and issuers.

We recommend that provider groups be clearly defined as a combination of TINs with their associated NPIs that collectively contract rates with the insurer. Specifically, the *negotiated price object* could allow multiple rates for the same provider group based on defined provider characteristics, using standardized fields similar to the existing service code array to indicate when different provider types are paid different amounts (with a blank value indicating uniform payment across provider types). In parallel, the *provider reference object* could be expanded to convert each NPI into a structured *NPI object* that includes attributes such as specialty, credentials, and practice locations, which could then be referenced by the *negotiated price object* to indicate when prices vary based on specialty. CMS should also seek broader stakeholder input on common real-world approaches to price categories—such as by credential, specialty, or location—to ensure schema updates reflect typical contracting practices and improve consistency across issuers.

Expansion of scope

The Departments should encourage expansion of the reporting requirements to additional plan types, including Medicare Advantage and Medicaid, using appropriate statutory and regulatory authority. Specifically, the Departments have the authority to require price information to be shared by other insurers and should leverage this authority to require Medicare Advantage and Managed Medicaid plans to report relevant pricing information.

Output formats and 'rectangular' datasets

We recommend that CMS require support for standardized rectangular output formats for TiC data to improve usability for analysts and reduce processing burden. Rectangular, table-based data structures are the standard format used across analytics, research, and data science workflows. Requiring rectangular outputs would allow users to work directly with TiC data using common analytic tools rather than requiring complex transformation steps.

In our recent publication, we explore this recommendation in depth and argue that rectangular formats align with how TiC data are already stored internally by issuers. Currently, issuers serialize relational database data into nested JSON structures for publication, requiring downstream users to reconstruct

relational tables before performing analysis. This creates unnecessary processing overhead and increases the risk of analytic errors without improving data utility.

While JSON is widely used for data exchange and can support Extract, Transform, and Load (ETL) workflows, it is not optimized for large-scale analytic consumption. Working with TiC data files in JSON format typically requires parsing entire files, reconstructing provider reference tables, and rebuilding relational joins before analysis can begin. In practice, the TiC JSON schema functions as a relational database spread across multiple nested structures, but without the performance or usability benefits of a true relational format.

We recognize the Department's concern that flat CSV files can become difficult to work with if they are fully flattened and contain repeated values across millions of rows. However, this concern can be addressed by requiring relational rectangular outputs consisting of multiple linked tables rather than single flattened files. Modern tools such as Python or DuckDB allow even modest personal computers to efficiently query large relational CSV or Parquet datasets, particularly when data are properly normalized.

Empirical testing demonstrates that rectangular formats can substantially reduce file size compared to JSON. In our analysis of 2,249 TiC files, relational CSV files were 47% smaller than the JSON files. In addition, the relational format makes them much faster to work with.

For these reasons, **we recommend that the Departments require the use of a standardized relational rectangular database structure**. CMS should define a canonical schema and allow issuers to publish data in multiple rectangular formats. While CSV is a standardized format that is well used, the Apache Parquet format may be an even better option as it is both standardized and significantly smaller and faster to work with. **We recommend that the Departments not define the file type as CSV but allow a full review of CSV and Parquet files before one is chosen as the final format.**

There is a strong federal precedent for this approach. CMS already distributes hospital cost report data using linked relational tables and has supported high-performance analytic formats (including Parquet) in other public data releases. Requiring relational rectangular outputs for TiC data would be consistent with existing CMS data publication practices and would future-proof the regulation as data infrastructure evolves.

Adopting relational rectangular output standards would improve accessibility, reduce insurer and user processing burden, improve analytic reproducibility, and expand the number of organizations capable of using TiC data. These changes align directly with CMS's goal of making machine-readable files more usable, standardized, and analytically meaningful.

Information to include for each plan

To ensure that TiC data can be meaningfully interpreted and used by regulators, employers, researchers, and other stakeholders, the Departments should require a standardized and comprehensive set of plan-level information to be included in the Table of Contents files. At a minimum, this information should identify the name of the employer sponsoring the plan and a corresponding employer identifier, the plan name, and a link to publicly available plan documentation that describes covered benefits and cost-sharing structures. Including total enrollment at the plan level is essential for contextualizing pricing information and understanding the relative scale and market impact of individual plans.

The Table of Contents file should also explicitly identify all provider networks associated with each plan, recognizing that many self-insured plans rely on shared or overlapping networks rather than bespoke provider arrangements. For each referenced network, plans should report total enrollment at the network level and provide a link to network-specific resources, such as a provider directory or searchable network tool. Together, these elements would allow users to accurately connect negotiated rates to the plans and networks under which they apply, improve the ability to aggregate and compare prices across plans and markets, and reduce ambiguity that currently limits the real-world utility of TiC data.

While more difficult to assemble, **it would be even more advantageous to include county-level counts of enrollment**. This would help analysts, regulators and business leaders better understand and act on enrollment. Similar methodologies that are used when releasing data on county-level enrollment in Medicare Advantage plans could be followed.

Additional proposed-rule topics to address

We support the Department's proposed changes to the Out-of-network Allowed Amount File. Lowering the inclusion threshold for number of claims per item from 20 to 11, increasing the lookback period from 180 days to nine months, and changing the reporting period to six months will result in more usable information and improve the overall quality of data in the file.

We also support changing the update frequency to a quarterly cadence as there isn't sufficient change from month to month to require a monthly update. A quarterly update cadence will also reduce the required effort for the insurers and free up resources for other TiC data release related tasks.

Including enrollment volume at the plan level will be a positive addition to the TiC and will increase its usability. For example, plan level enrollment data can be used to understand how total enrollment at

the plan level relates to prices. This is important since bigger plans might be in a better position to renegotiate and lower prices.

The implementation timeline of the proposed rules should be moved up and the plans should be required to be in compliance by July 1, 2027.

The proposals outlined by the Departments to improve and standardize the TiC data and make it easier to access and use should also be considered for the release of prescription drug prices in a machine-readable file format.