



Handl Health

■ SEPTEMBER 1, 2026

**How confident are you
in your healthcare
pricing data?**

www.handlhealth.com/

■ Executive Summary

It was to be a new era of healthcare price transparency. Before getting care, patients would be able to research rates for procedures (think airline price aggregators) and then book an appointment with the provider promising the best service for the price. This would drive down healthcare prices, focusing patients on low-cost, high-value providers.

That is what the Centers for Medicare and Medicaid Services envisioned when it first published **Transparency in Coverage (TiC) rules in 2020**. The reality is completely different. The federally mandated Machine-readable price transparency files have historically been filled with rates that were, at best, hard to use. And impossible for consumer consumption. The regulations unleashed a variety of procedure rate types that didn't compare cleanly, meaning that validation with data set upon data set became a crucial part of the process. The same carrier published the same network under multiple file structures across regions. Some providers made their fees inscrutable, if they published them at all. (Provider regulation compliance is an ongoing issue).

And that's not even considering the sheer magnitude of the files:



A single MRF could clock in at 50 gigabytes, making a national rate pull into a multi-terabyte endeavor. Computer processing power became a significant barrier to entry.

By themselves, healthcare price transparency rate files proved confusing at best and fairly useless as far as healthcare pricing analysis goes. Using the pricing information began with a huge data engineering challenge that, once solved, became a data interpretation challenge. In 2026, the quality is now better, but challenges still exist: **variability, incompleteness and ambiguity in MRF data.**

So how can you use price transparency data to lower the cost of care? Should employers, insurance brokers and other data users take price transparency information at face value?



Handl Health ingests billions of rates from payer and provider MRFs,

transforming and normalizing the data into searchable datasets for use by employers, insurance brokers and TPAs. To contextualize these challenges, and provide usable information that smooths out data inconsistencies, Handl Health uses proprietary confidence scoring. In this guide, you'll learn all about the evolution of that healthcare pricing data, the interpretation and enrichment of the files, and how Handl uses robust data to create analyses that are transforming the self-funded health insurance market.

■ What is an MRF?

Federal price transparency regulations mandate that carriers and providers publish pricing data in files designed for computer consumption (don't try to read one!). From a payer perspective, these files contain negotiated rates paid per procedure and allowed amounts for out-of-network care. For a hospital, MRF files contain gross charges, discounted cash rates and negotiated carrier charges for each medical identifier. Payers publish updated information every month, and providers are mandated to refresh their information annually.

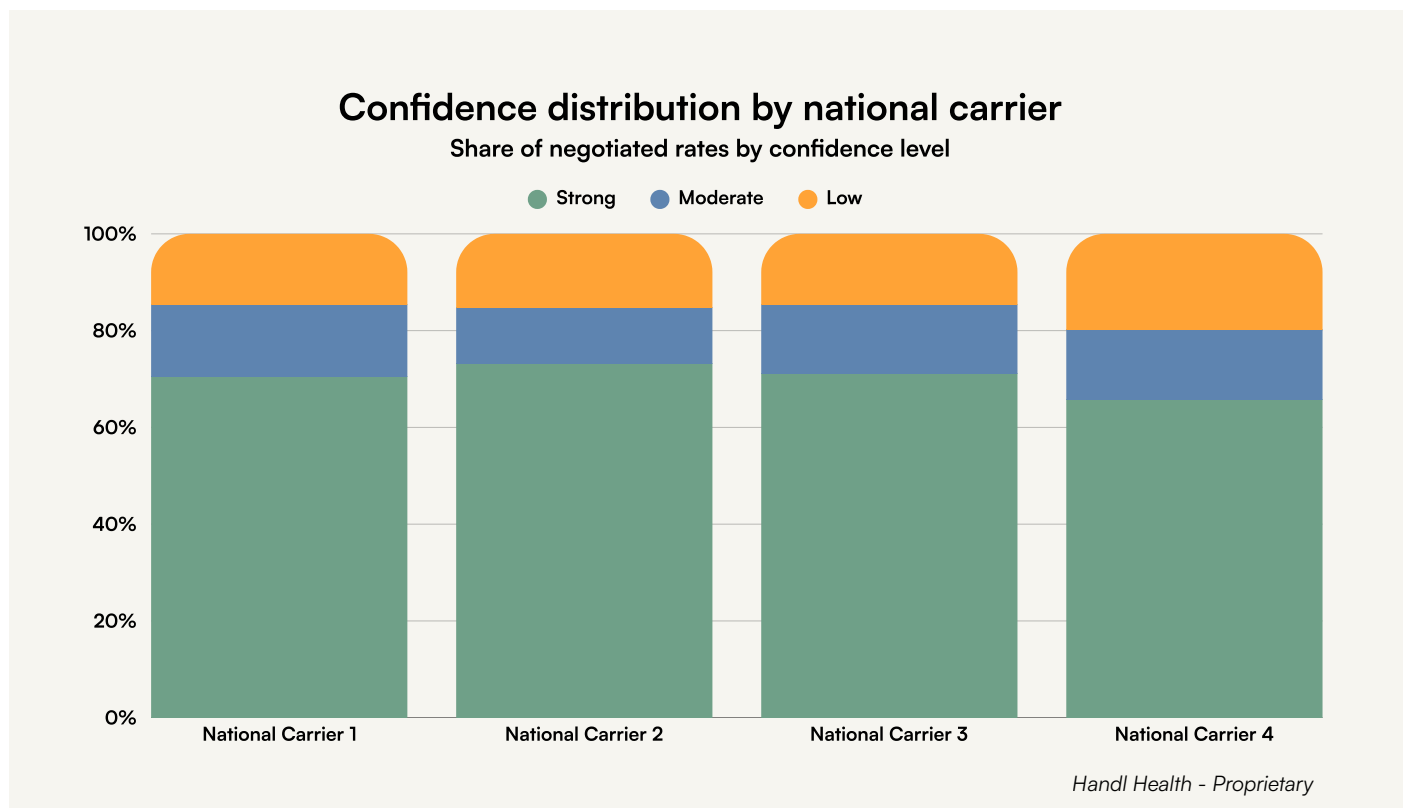
Every MRF varies in terms of data quality and fidelity based on the organization that's providing the data, the vendor being used to generate the file, and a host of other factors. But there's a general rule of thumb: the larger the entity, the more robust the transparency data. That's because large organizations can work with third parties to produce the files. What about smaller payers and providers? The data can get a little murky because it's simply not available, the vendor producing the MRF files isn't as thorough, or both.

When healthcare price transparency regulations first rolled out, the industry became hyper-focused on compliance metrics: Who published MRFs? How many files existed? Which carriers and providers were participating? At the time, that made sense:

**In the early years,
compliance was below**

20%

The industry believed that more published data would lead to a clearer understanding of healthcare pricing. Then technology companies actually started working with the data.



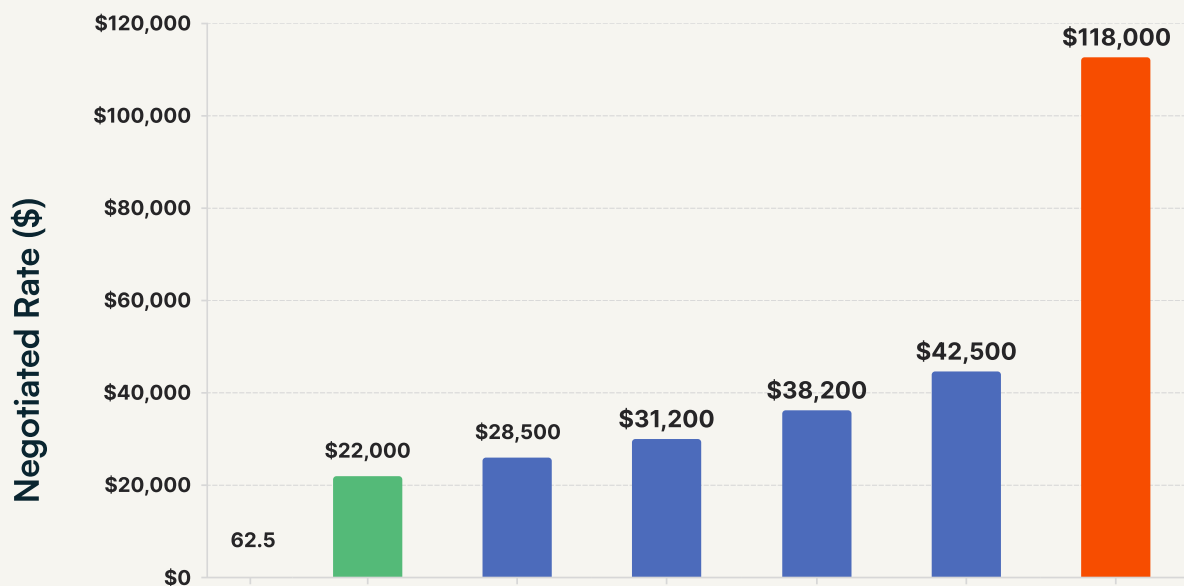
What became obvious very quickly is that compliance and usability are not the same thing. A file existing doesn't automatically make it useful. Some MRFs contain highly reliable negotiated rate data with strong provider and plan coverage. Others are sparse, inconsistent, structurally difficult to interpret, or are missing enough context that the output becomes misleading if used at face value.

MRFs are relatively new, but healthcare pricing analysis is an old science. Historically, insurance brokers searching for the best deal for their clients have relied on the Uniform Discount and Data Specification (UDS) data, jointly built by the carriers and the major consulting houses. Brokers used this claims information submitted by insurance carriers to analyze network discounts. The issue? Carriers only submit blanket discount data on a macro level — “Chicago providers have negotiated a 30% discount” — and there's a significant reporting delay. **Until the rise of price transparency data, UDS information was the only way** brokers could benchmark healthcare pricing, evaluate networks, reprice claims and create new plan designs. “Price by proxy” is a poor substitute for real-world pricing data.

■ What is a confidence score and why does it matter?

Price transparency data is messy. And inconsistent. It would not be smart to restrict hospital or carrier pricing analysis, network evaluation or any other report to MRF data as the only source of information. That's why Handl has developed data confidence scores: trust but verify, verify, verify.

To see the variety of MRF data, and why confidence scoring might be needed, take a look at the rates for a total knee arthroplasty, CPT 27447. That code returns rates from \$22,000 at an ambulatory surgery center to \$42,000 at an academic medical center. But there's also one rate for \$118,000, which is 8.3x the Medicare rate of \$14,200. Can these rates be trusted? What about that enormous one?



Handl Health - Proprietary

Confidence scores measure the reliability of a published negotiated rate compared to expected pricing behavior. When measuring confidence scores, Handl Health doesn't simply look at other MRFs to get a sense of pricing; that could easily skew our confidence and lead customers down the path of incorrect analyses. MRFs are used as a benchmark when measuring data confidence, but the more important data comes from historical claims information and acceptable Medicare thresholds. The confidence sits outside what someone might have published in an MRF.

To assess whether any rate is reasonable, it is compared against multiple independent benchmarks including Medicare fee schedules, commercial market data, plan-level averages and claims data. A z-score is then calculated relative to the benchmark distribution, measuring standard deviation from the mean. Rates that exceed a defined z-score threshold are flagged as statistical outliers and assigned lower confidence scores.



Medicare

Compare the MRF rate relative to expected Medicare reimbursement for the same code and setting, with acceptable deviation thresholds applied by service category. Handl Health digs into Inpatient Prospective Payment System (IPPS) and Outpatient Prospective Payment System (OPPS) rates published by the CMS on both a code and locality level. If the MRF rate exceeds the acceptable Medicare multiplier threshold, it could merit enhanced scrutiny.



Claims

Claims data reflects what payers have actually reimbursed providers for equivalent services, an ideal real-world benchmark. But since claims are by definition historic information, Handl Health applies a yearly trend factor to update the data, accounting for inflation, contract renegotiations, and broader market rate movements.

Once trended, claims-derived benchmarks are used to evaluate whether a given rate falls within an acceptable range. Outliers are assigned reduced confidence scores and excluded from downstream analysis.

Other MRFs



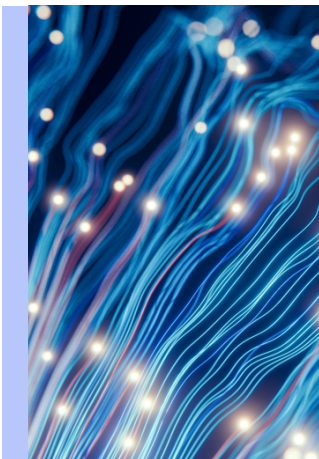
To ensure everything makes sense, the rate should be judged against pricing data published by similar payers. MRFs from similar payers are weighted more conservatively given the variability in how payers report and structure rates.

Every piece of data requires context. Medicare and claims benchmarks carry the most weight in this process because they reflect rates that have been directly observed and verified in practice.

Rates falling outside acceptable deviation bands are flagged as outliers and excluded from Handl analyses, giving users confidence in Handl data. Healthcare pricing analysis, network studies and care steerage reports all start with MRFs but quickly evolve from there. After all, the best reports start with the best healthcare data.

Confidence scores are assigned to data so the system can prioritize the most reliable rates during integration; flag potential data quality issues; and provide a quantifiable indicator of data reliability.

In the knee replacement example, a \$118,000 rate for the procedure is a clear statistical outlier. It's not representative of what is typically charged, so the rate is excluded. Chalk it up to another case of bad data. **But the four hospitals offering the same exact procedure with a \$14,000 spread? That's just the reality of healthcare pricing.**



Want some AI with your MRFs?

In the healthcare industry, human touchpoints are crucial, so AI should be used as an augmentation tool and not as a replacement for health plan management. AI is extremely useful in MRF collection, as agents to traverse hospital websites looking for files, but AI should not be used for sensitive data processing. At Handl Health, AI enables clients to go deeper with analyses and better construct truly custom, long-term health plans.

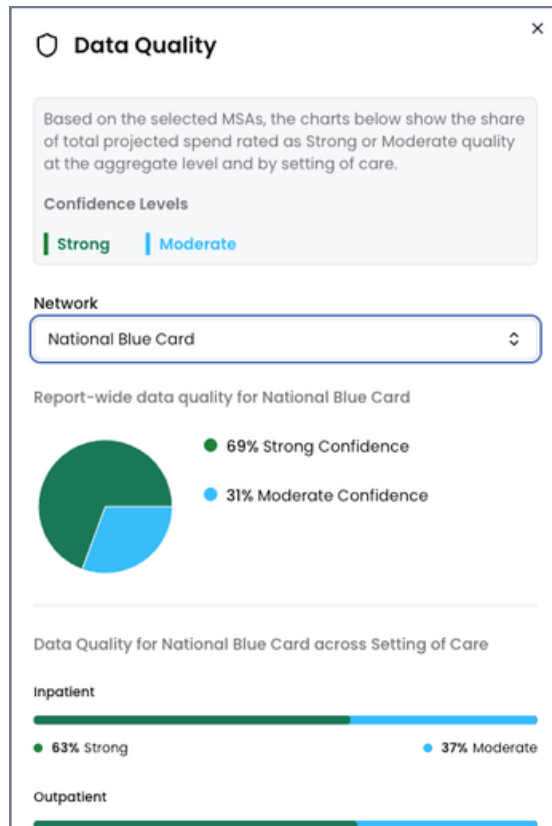
■ Putting confidence scores into context

After measuring how far the published rate deviates from benchmarks, Handl assigns a confidence score: Strong, Moderate and Low. These weighted synthesis confidence scores are between 0 and 100; higher values mean stronger alignment across benchmarks. A moderate score is not necessarily bad and should be used in analysis.

Rates that fall significantly outside of expected pricing patterns are flagged as outliers. This helps identify issues like extreme under- or over-reporting by hospitals or carriers; partial file updates or missing modifiers; and misalignment between care setting and billing class. These outliers remain in the dataset for transparency but are clearly flagged to avoid contaminating analytics or benchmarking.

Tier		Treatment
	Strong	Ready for analysis!
	Moderate	Useful in analysis but flagged. Imputation may supplement depending on category quality coverage.
	Low	Excluded from direct use.

■ Derived Data: What happens when MRFs aren't complete?



MRF data has gaps. When confidence scores are low, Handl tries to fill these gaps with other data, combining carrier-specific pricing patterns with market benchmarks to produce reliable estimates that support better decision-making. This is derived data.

Instead of using generic benchmarks, **Handl adjusts estimates based on actual pricing patterns in the market.** This means the derived, imputed data reflects real differences in how carriers price services—like regional variation or negotiated rates—enabling accurate analysis grounded in reality.

By thoughtfully estimating missing rates, Handl completes the data landscape, giving users access to an unrivaled amount of data.

Users are empowered to compare costs across networks and providers with greater confidence and clarity.

Imputation, which leads to derived data, is a distinct process from confidence scoring and rate selection. **It addresses a specific gap: billing codes that appear in an employer's claims spend but have no published rate in any MRF** at the MSA and plan level. Excluding these codes from analysis would understate projected costs. That's why Handl derives a rate using patterns observed in the source MRF data and claims benchmarks.

The unit of analysis is the clinical care category, a grouping of related billing codes representing a coherent service type (inpatient orthopedic surgical or outpatient imaging for example). Within each category, Handl calculates the share of weighted spend covered by high-quality MRF rates. This quality coverage metric drives imputation along a linear spectrum between the carrier differential and claims benchmarks.

Anchor		Description
	Carrier Differential	Imputed rates are derived by extending the carrier-specific pricing differential observed in the available MRF data for that category. This preserves how the carrier prices this type of care relative to the market.
	Claims Anchor	Imputed rates are anchored to claims-based benchmarks for the relevant geography and care setting, providing a market-grounded estimate when MRF patterns are too sparse to characterize carrier behavior.

As quality coverage within a category increases, imputed rates shift proportionally toward the carrier differential; as it decreases, they shift toward the claims anchor. This means the methodology self-adjusts to the available evidence rather than applying a fixed rule.

All imputed rates are labeled as derived data. The data coverage breakdown in each Handl Health report shows exactly which portions of the analysis rely on direct MRF rates versus derived estimates. But derived data is not a bad thing. Without imputed rates, many brokers could not effectively compare networks with spotty data availability. MRF-based analysis would also be limited to only those providers and services where carriers have published complete files, excluding the most complex and highest-cost cases where pricing visibility matters most.

■ How Handl uses MRF pricing data

So when multiple parties report different prices on the same procedure, what happens? Handl integrates carrier and hospital MRFs at the plan level. Rates are matched across both sources at a common set of data fields:

- **National Provider Identifier (NPI)**
- **Tax Identification Number (TIN)**
- **Billing code:** CPT, DRG or applicable code type
- **Billing class:** institutional vs. professional
- **Other data fields as published:** rate type, contract type, unit, etc.

Any rate that maps to the same combination is treated as a candidate rate for that slot. When multiple candidates exist, the highest-confidence rate wins. This applies whether the conflict is across sources or within a single source. Source preference is not assumed; it is determined by the confidence score.

Rate selection is the output of cross-validation. Once candidates for a given combination have been scored, the highest-confidence rate is selected as the operative rate for analysis.



Single Candidate

Used directly in analysis.



Multiple Candidates

The rate with the strongest alignment across Medicare, claims and MRF peer benchmarks wins. The methodology is algorithmic but not binary. Similar rates are flagged for review.

■ MRFs are the foundation of Handl data

Handl data draws on multiple sources beginning, but not ending, with price transparency MRFs.

Carrier Transparency in Coverage MRFs:

This shows what different carriers have negotiated with the same provider for the same service.

Hospital and provider MRFs:

This dataset is cross-validated against carrier files to eliminate duplicative rates; the datapoint with the highest confidence score remains.

Commercial claims utilization patterns:

Used for contextual benchmarking.

Provider Directory:

Handl can enrich MRF data with provider directories straight from carrier feeds.

HANDL HEALTH TYPICALLY FLAGS BETWEEN

20% to 30%

OF MRF RATES BECAUSE THEY DON'T MEET THRESHOLDS FOR USABLE DATA.

FOR THE NATIONAL CARRIERS, HANDL TYPICALLY REJECTS ANYWHERE FROM

21% to 27%

OF THE DATA IT RECEIVES

Handl routinely comes across data that is compliant with the law on its face but can't be used in any analysis. Take a look at the first half of 2026:

Handl fielded 47 requests to add networks into its data set, and 11 did not meet data requirements.

Handl couldn't use this information because **the data was missing, unstable or it just wasn't strong enough to provide confidence** that the rates were accurate.

The government is currently applying pressure to make sure providers are making their pricing data available, which is a positive development. But the more important conversation isn't whether a file was published. It's whether the data is reliable enough to support real decision-making. Instead of compliance numbers, we need to look at how much MRF data is actually being used and how technology providers use derived data to present a robust picture of healthcare industry pricing. All of this is hard to enforce because it takes a lot of work to "look under the hood" on MRF rates and determine if MRF files are actually compliant or just box checking.

Step by step

From MRF data to insight

Clean

Remove inconsistencies, duplicates, formatting errors, implausibly high/low rates and ghost rates. Ghost rates are data for providers or services with no evidence of active delivery.

Collect

Ingest available carrier and hospital files, normalized to a common schema.

Weigh

Billing code: Utilization-per-1,000 weights based on commercial claims data representing approximately 98% of a typical employer's medical spend across roughly 5,000 billing codes. This ensures that high-volume, high-spend services drive the analysis rather than rarely used codes.

Provider market share: Rates are weighted by expected utilization for a given service and provider market share. This gives providers that serve a larger share of the relevant population a greater influence on network cost projections.

Aggregate: Structuring weighted rates by region (state, MSA), specialty, service category, place of service, and network.

Validate

Cross-validation against commercial claims data and Medicare fee schedules to identify rates that fall outside expected ranges. Rates are compared to claims-derived paid amounts and CMS fee-for-service baselines to ensure MRF data is directionally sound and decision-ready.

■ Confidence scores in the real world

Confidence scores are meant to help users find the strongest data for their analyses. In a large metro area for a specific product type (a PPO or HMO for one of the large insurance carriers), data confidence is high. That has to do with the amount and breadth of the data available. A smaller population, maybe a micro MSA for a regional plan, will have to rely much more on imputed rates. Don't shy away from this derived data. Remember though, it's important to consider when to use the confidence scores as directional guidance versus mere recommendations.

MRF data, both imputed data and information with high confidence, might lead brokers to different conclusions from their UDS datasets. Where UDS expresses performance as a discount off billed charges using claims that can be 2-3 years old, Handl Health shows current contracted rates in absolute dollars and as a percent of Medicare.

Together, they validate or challenge a carrier's self-reported savings narrative with objective, external data.



MRF data replaces the need for UDS in the majority of real-world scenarios: outside the marketing window (UDS has strict data-use limitations restricting it to active carrier negotiations); when no NDA or data sharing agreement is in place; when time is a constraint (the Handl platform produces results in minutes versus weeks for UDS);

or when client-ready output is needed directly from the analysis. Because the Handl platform requires no setup agreements and uses MRF data with a 90-day lookback, **consultants can run network evaluations with timely data year-round**, not just during RFP cycles. So when UDS and MRF-fueled (and even derived) data are in conflict, the better choice is usually price transparency data, not out-of-date claims information.

About Handl Health

Handl makes it possible for every employer to offer affordable, personalized healthcare designed around the unique needs of their people. By empowering insurance brokers, third-party administrators and carriers to evaluate, design and monitor health plans, Handl is promoting healthcare that delivers the right care at the right time and at a price working families can afford.

A screenshot of a Handl Health interface on a dark blue background. A white card displays the following information:
Procedure: **Colonoscopy**
Location: Hospital A • Doctor: Dr. Adams
Table of costs:

Total cost before coverage	\$3,292
What your plan pays	\$1,793
What you pay	\$1,499

A "Continue" button is at the bottom.

Parallel intelligence for employer-sponsored health plans.

 www.handlhealth.com

