# MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12-05_QCP_in-network-rates.json.gz

Size: 66.20 MB • MD5: ba7f773c2c22beea24b9627fa2b0b9ee

Payer: **Quality Care Partners** • File Date: **12/4/2025** • Generated: **2026-04-24 20:32 EDT** • Tool Version: **1.0.0** • Elapsed: **468.50s**

**74.2**

**Limited Reliability**

Errors: 1 • Warnings: 11 • Info: 3

## TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12_plan_ref_98d7022b7d7165738b14c89261e12b3b_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12-10_QCP_in-network-rates.json.gz

Mirror: s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12-10_QCP_in-network-rates.json.gz

| Plan Name | Plan ID | Issuer / Sponsor | Market |
|---|---|---|---|
| QCP | 0123456789 (custom) | — | — |

## CMS Official Schema Validation

**PASSED** — File conforms to the CMS in-network-rates schema.

## Dimension Scores

| Dimension | Score | Weight | Findings |
|---|---|---|---|
| Schema Integrity | 79.4 | 30% | 4 |
| Provider Mapping | 99.9 | 15% | 4 |
| Code Coverage | 70.0 | 15% | 2 |
| Pricing Sanity | 62.1 | 40% | 5 |

# Schema Integrity — Findings

Score: 79.4

**WARNING** `header_conditional`

last_updated_on '12/4/2025' is not a valid YYYY-MM-DD date

**WARNING** `header_conditional`

Partial plan fields present — missing: ['plan_market_type']

**WARNING** `rates_without_providers`

340256 negotiated_rate entries have neither provider_groups nor provider_references

**INFO** `far_future_expiration`

4103276 expiration_dates are more than 3 years in the future

# Provider Mapping — Findings

Score: 99.9

**WARNING** `npi_validity`

0.04% of NPIs failed Luhn checksum validation (2 of 5503)

- 1588679012
- 1810341789

---

**WARNING** `ein_validity`

0.31% of EINs failed IRS prefix validation (18 of 5748)

- 281723399
- 282504837
- 282884410
- 283428410
- 286429330
- *… and 5 more*

---

**WARNING** `empty_npi_groups`

245 provider groups contain no NPIs

---

**INFO** `duplicate_npis`

325 NPIs appear in more than one provider group

---

# Code Coverage — Findings

Score: 70.0

**WARNING** `billing_code_format`

156736 HCPCS codes do not match expected format

**WARNING** `duplicate_billing_codes`

14702 billing codes appear in more than one in_network item (99.1%)

# Pricing Sanity — Findings

Score: 62.1

**INFO** `percentage_rates`

495 percentage rates (0.0%) — values represent % of a reference rate, not dollar amounts; excluded from spread analysis

**ERROR** `zero_rates`

101642 zero-dollar rates (2.48%) — CMS schema requires negotiated_rate > 0 (exclusiveMinimum)

**WARNING** `rate_spread_by_class`

billing_class='institutional' / negotiated_type='fee schedule': P95/P50 spread is 19.3x (threshold: 10x, N=2,096,094 (1,000 sampled), high confidence)

**WARNING** `rate_spread_by_class`

billing_class='professional' / negotiated_type='fee schedule': P95/P50 spread is 14.4x (threshold: 5x, N=2,006,681 (1,000 sampled), high confidence)

**WARNING** `per_code_rate_spread`

1057 rate contexts have a max/min ratio exceeding the type-specific threshold (20x professional / 50x facility, min 3 occurrences required). Each context is a unique combination of all 10 rate-key dimensions. n= shows how many distinct provider rates exist for that exact context.

| Code | Code Type | Neg. Type | Billing Class | Arrangement | Setting | Min | Median | Mean | Max | Ratio | n |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Q0249** | HCPCS | fee schedule | institutional | ffs | — | $7.57 | $3293.76 | $4462.93 | $14129.80 | **1866.8×** | 223 |
| **92953** | CPT | fee schedule | institutional | ffs | — | $0.81 | $1.66 | $409.00 | $1435.83 | **1772.6×** | 167 |
| **Q0249** | HCPCS | fee schedule | professional | ffs | — | $6.81 | $6572.00 | $3561.93 | $11829.60 | **1736.6×** | 220 |
| **95004** | CPT | fee schedule | institutional | ffs | — | $3.38 | $6.93 | $638.47 | $2232.34 | **660.5×** | 167 |
| **95044** | CPT | fee schedule | institutional | ffs | — | $4.66 | $9.55 | $640.02 | $2232.34 | **479.0×** | 167 |
| **86905** | CPT | fee schedule | institutional | ffs | — | $3.64 | $7.47 | $226.40 | $789.29 | **216.9×** | 168 |
| **63685** | CPT | fee schedule | institutional | ffs | — | $315.24 | $646.24 | $19383.65 | $66865.00 | **212.1×** | 167 |
| **33289** | CPT | fee schedule | institutional | ffs | — | $296.20 | $607.21 | $18085.27 | $62377.31 | **210.6×** | 167 |
| **33231** | CPT | fee schedule | institutional | ffs | — | $360.56 | $739.15 | $20428.82 | $70349.35 | **195.1×** | 167 |
| **33264** | CPT | fee schedule | institutional | ffs | — | $362.67 | $743.47 | $20431.38 | $70349.35 | **194.0×** | 167 |

# Recommended Actions

1. **pricing** `zero_rates` **P1**

101642 zero-dollar rates (2.48%) — CMS schema requires negotiated_rate > 0 (exclusiveMinimum)

2. provider_mapping npi_validity P2

0.04% of NPIs failed Luhn checksum validation (2 of 5503)

3. provider_mapping ein_validity P2

0.31% of EINs failed IRS prefix validation (18 of 5748)

4. provider_mapping empty_npi_groups P2

245 provider groups contain no NPIs

5. pricing rate_spread_by_class P2

billing_class='institutional' / negotiated_type='fee schedule': P95/P50 spread is 19.3x (threshold: 10x, N=2,096,094 (1,000 sampled), high confidence)

6. pricing rate_spread_by_class P2

billing_class='professional' / negotiated_type='fee schedule': P95/P50 spread is 14.4x (threshold: 5x, N=2,006,681 (1,000 sampled), high confidence)

7. pricing per_code_rate_spread P2

1057 rate contexts have a max/min ratio exceeding the type-specific threshold (20x professional / 50x facility, min 3 occurrences required). Each context is a unique combination of all 10 rate-key dimensions. n= shows how many distinct provider rates exist for that exact context.

8. schema header_conditional P2

last_updated_on '12/4/2025' is not a valid YYYY-MM-DD date

9. schema header_conditional P2

Partial plan fields present — missing: ['plan_market_type']

10. **schema** `rates_without_providers` **P2**

340256 negotiated_rate entries have neither provider_groups nor provider_references

11. **code_coverage** `billing_code_format` **P3**

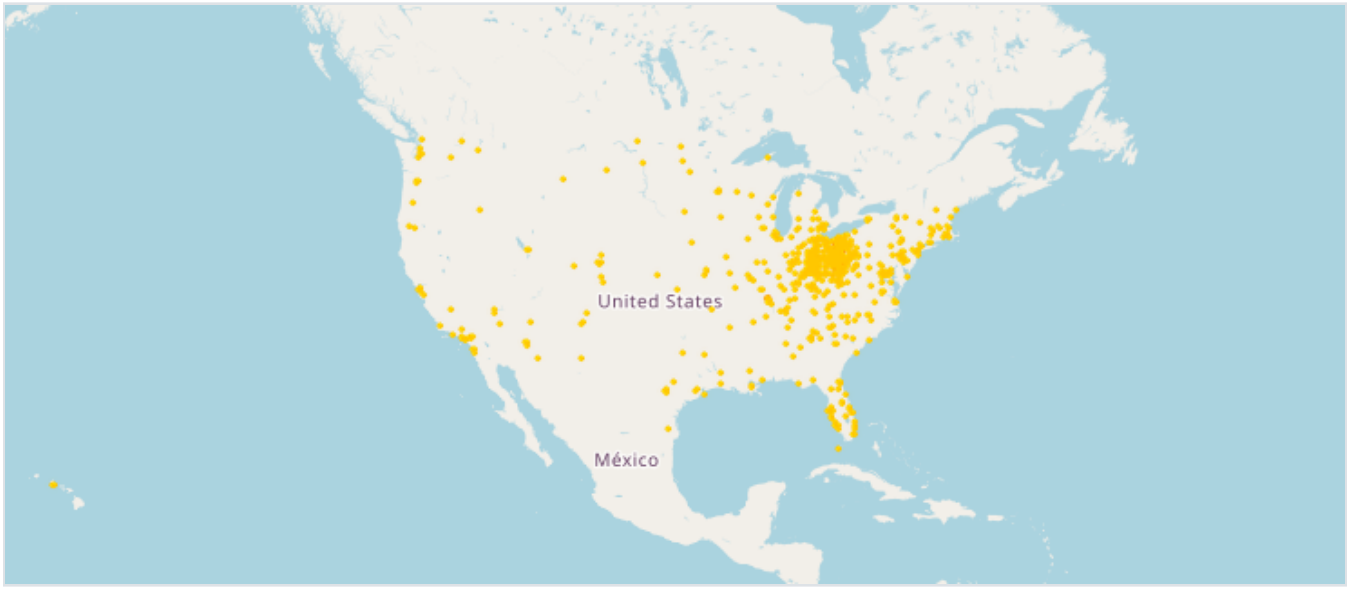
156736 HCPCS codes do not match expected format

12. **code_coverage** `duplicate_billing_codes` **P3**

14702 billing codes appear in more than one in_network item (99.1%)

## Provider Geographic Coverage

5149 unique NPIs found — 5139 geocoded (100%) — 864 zip codes represented.




## Schema Integrity — Metrics

| | |
|---|---|
| **header_missing_fields** | |
| **header_conditional_issues** | last_updated_on '12/4/2025' is not a valid YYYY-MM-DD date, Partial plan fields present — missing: ['plan_market_type'] |
| **file_age_days** | None |
| **items_total** | 4101037 |

| | |
|---|---|
| items_missing_required_pct | 0.0 |
| items_empty_rates | 0 |
| prices_total | 4103276 |
| prices_missing_required_pct | 0.0 |
| prices_missing_field_breakdown | |
| prices_missing_service_code | 0 |
| prices_invalid_billing_class | 0 |
| rates_without_providers | 340256 |
| negotiation_arrangements | ffs: 4101037 |
| billing_code_types | CSTM-ALL: 922<br>HCPCS: 1215810<br>CPT: 2884111<br>RC: 194 |
| expired_prices | 0 |
| invalid_expiration_format | 0 |

## Provider Mapping — Metrics

| | |
|---|---|
| provider_references_in_file | 5748 |
| provider_group_ids_referenced | 5748 |
| unresolved_references | 0 |
| resolution_rate_pct | 100.0 |
| npis_validated | 5503 |
| invalid_npi_count | 2 |
| npi_validity_rate_pct | 99.96 |
| invalid_npi_examples | 1588679012, 1810341789 |
| eins_validated | 5748 |
| invalid_ein_count | 18 |
| ein_validity_rate_pct | 99.69 |
| invalid_ein_examples | 281723399, 282504837, 282884410, 283428410, 283428410, 283428410, 286429330, 287641240, 287743571, 287841552 |

| | |
|---|---|
| empty_npi_groups | 245 |
| groups_without_tin | 0 |
| npis_in_multiple_groups | 325 |

## Code Coverage — Metrics

| | | |
|---|---|---|
| unique_codes_total | 14838 | |
| duplicate_codes | 14702 | |
| duplicate_pct | 99.08 | |
| by_code_type | CSTM-ALL | 396 |
| | HCPCS | 5012 |
| | CPT | 9383 |
| | RC | 47 |
| unknown_code_types | | |
| format_invalid_by_type | HCPCS | 156736 |
| codes_not_in_reference | reference_not_loaded | |

| | |
|---|---|
| most_frequent_codes | |

| Type | Code | Occurrences |
|---|---|---|
| HCPCS | E0950 | 1,782 |
| HCPCS | E0951 | 1,782 |
| HCPCS | E0952 | 1,782 |
| HCPCS | E0953 | 1,782 |
| HCPCS | E0954 | 1,782 |
| HCPCS | E0956 | 1,782 |
| HCPCS | E0957 | 1,782 |
| HCPCS | E0960 | 1,782 |
| HCPCS | E0973 | 1,782 |
| HCPCS | E0978 | 1,782 |
| HCPCS | E0981 | 1,782 |
| HCPCS | E0982 | 1,782 |
| HCPCS | E0990 | 1,782 |
| HCPCS | E0995 | 1,782 |
| HCPCS | E1016 | 1,782 |
| HCPCS | E2208 | 1,782 |
| HCPCS | E2209 | 1,782 |
| HCPCS | E2210 | 1,782 |
| HCPCS | E2323 | 1,782 |
| HCPCS | E2324 | 1,782 |

## Pricing Sanity — Metrics

| | |
|---|---|
| total_prices_checked | 4103276 |
| total_rates | 4102781 |
| per_diem_rates | 0 |
| percentage_rates | 495 |
| negative_rates | 0 |
| zero_rates | 101642 |
| extreme_high_rates | 2516 |
| extreme_low_rates | 914 |

**rate_distribution**

| | |
|---|---|
| **sample_n** | 4102781 |
| **sample_k** | 5000 |
| **confidence** | high |
| **p5** | 2.3878125000000003 |
| **p25** | 47.47275 |
| **p50** | 254.66295000000002 |
| **p75** | 1025.4792 |
| **p95** | 4816.215 |
| **p99** | 14216.611499999999 |

**by_billing_class**

| Class / Type | Count | Median | p25 | p75 | p95 | Confidence |
|---|---|---|---|---|---|---|
| **institutional/ fee schedule** | 2,096,094 | 383.3 | 85.3 | 1469.6 | 7387.3 | high |
| **professional/ fee schedule** | 2,006,681 | 146.4 | 36.7 | 647.7 | 2111.1 | high |
| **institutional/ derived** | 3 | 1.0 | 1.0 | 1.0 | 1.0 | low |
| **professional/ derived** | 3 | 1.0 | 1.0 | 1.0 | 1.0 | low |

**negotiated_types**

| | |
|---|---|
| **fee schedule** | 4102775 |
| **derived** | 6 |

**unique_rate_contexts** 34419

**rate_key_dimension_validity**

| | |
|---|---|
| **invalid_negotiated_type** | 0 |
| **invalid_negotiated_types_seen** | {} |
| **invalid_setting** | 0 |
| **invalid_settings_seen** | {} |
| **invalid_severity_of_illness** | 0 |
| **severity_on_non_apr_drg** | 0 |
| **institutional_with_service_codes** | 0 |
| **invalid_service_code_format** | 0 |
| **billing_code_modifier_too_long** | 0 |

# Scoring Methodology

Embedded in this report at generation time.

## Overall Score

Weighted sum of four structural dimensions, normalized to a 0–100 scale.

| | | |
|---|---|---|
| **Normalized Weights** | **Schema Integrity** | 30% |
| | **Provider Mapping** | 15% |
| | **Code Coverage** | 15% |
| | **Pricing Sanity** | 40% |
| **Confidence Bands** | **High** | ≥90 |
| | **Usable With Caution** | ≥75 |
| | **Limited Reliability** | ≥60 |
| | **Not Usable** | <60 |
| **Score Caps** | **Raw Json Errors Only → 74.0** | Native JSON syntax errors in the unpatched source file. File must be re-exported by the payer; scoring reflects auto-patched data only. |
| | **Cms Validation Failure Only → 65.0** | CMS official schema validator reports the file does not conform to the TIC spec. |
| | **Both Raw Json Errors And Cms Failure → 59.0** | Both native JSON syntax errors and CMS schema validation failure present. |

### Rate Context Key — 14-Tuple Field Coverage

Every rate in a CMS TIC MRF file is described by a 14-field tuple. Fields 1–10 form the rate-context key used to group and compare rates across the system. Fields 11–12 (provider, expiration date) are validated separately and excluded from the grouping key for analytical reasons. Each of the four scoring dimensions validates a distinct slice of this tuple — together they cover all 14 fields.

| Field | Validated by |
|---|---|
| `1 billing_code_type` | Schema (required field) + Code Coverage (enum + format validation) |
| `2 billing_code_type_version` | Schema (required field) |
| `3 billing_code` | Schema (required field) + Code Coverage (format, duplicates, reference lookup) |
| `4 billing_code_modifier` | Pricing (modifier length, key normalization) |
| `5 service_code` | Pricing (POS format, normalization, institutional-class check) |
| `6 negotiated_type` | Pricing (CMS TIC enum validation) |
| `7 billing_class` | Schema (CMS TIC enum validation) + Pricing (spread thresholds) |
| `8 negotiation_arrangement` | Schema (CMS TIC enum validation) + Pricing (FFS vs bundle/capitation gating) |
| `9 severity_of_illness` | Pricing (APR-DRG only, valid values 1–4) |

| Field | Validated by |
|---|---|
| `10 setting` | Pricing (CMS TIC enum validation) |
| `11 provider (NPI/EIN)` | Provider Mapping (Luhn checksum, IRS prefix, group resolution) — excluded from grouping key |
| `12 expiration_date` | Schema (date validity, far-future sanity) — excluded from grouping key |
| `13 additional_generic_notes` | not validated (free-text) |
| `14 negotiated_rate` | Pricing (negative/zero/extreme-value checks, spread analysis) |

- Fields 1–10 are the grouping key. Each unique combination is a distinct rate context — rates with different modifiers, POS codes, or arrangements land in separate buckets and are never compared against each other.
- Provider (field 11) is excluded from the key: the spread check is cross-provider by design. Partitioning by provider produces singleton buckets and eliminates the spread signal.
- Expiration date (field 12) is excluded because it is a contract lifecycle attribute, not a clinical context. Rates for the same service should be comparable regardless of when they expire.
- service_code (field 5) arrays are flattened and normalized before keying: '1' → '01', and a rate with ['11','22'] contributes to both the '11' and '22' buckets so rates are compared apples-to-apples by place of service.

## Schema Integrity

Validates required fields, enum values, conditional requirements, and date validity per the CMS TIC in-network-rates schema. Also checks file freshness and expiration date sanity.

**Method:** Penalty-based deductions from 100, capped per category.

| | |
|---|---|
| **per_missing_required_header_field** | 5 |
| **per_header_conditional_issue** | 2 |
| **freshness_warn** | 5 |
| **freshness_error** | 10 |
| **item_missing_fields_pct** | ×5 (cap 30) |
| **empty_rates_pct** | ×0.5 (cap 5) |
| **price_missing_fields_pct** | ×10 (cap 30) |
| **rates_without_providers_rate** | ×200 (cap 20) |
| **expired_prices_pct** | ×0.5 (cap 5) |

| | |
|---|---|
| **file freshness warn days** | 45 |
| **file freshness error days** | 90 |
| **expiry far future years** | 3 |

## Provider Mapping

Verifies that all provider_group_id references in in_network items resolve to an entry in the provider_references array. Validates NPI integrity via Luhn checksum and EIN integrity via IRS-issued 2-digit prefix.

**Method:** Weighted component sum (not purely penalty-based).

| | |
|---|---|
| **provider_resolution (60%)** | `resolution_rate% × 0.60` |
| **npi_validity (30%)** | `(100 − invalid_npi_pct × 5) × 0.30` |
| **ein_validity (10%)** | `10 − (invalid_ein_pct × 0.1) [0% invalid → 10 pts, 100% invalid → 0 pts, linear]` |

## Code Coverage

Tracks every (billing_code_type, billing_code) pair and flags unrecognized CMS TIC code types, format violations for CPT/HCPCS/NDC, and duplicates (same code appearing in multiple in_network items).

**Method:** Penalty-based deductions from 100.

| | |
|---|---|
| **per_unknown_code_type** | `3 pts each (cap 20)` |
| **format_invalid_pct** | `× 0.5 (cap 10)` |
| **duplicate_code_pct** | `× 2 (cap 20)` |
| **codes_not_in_reference_pct** | `× 0.5 (cap 30) — only when reference set is loaded` |

## Pricing Sanity

Detects invalid rates (negative, zero, extreme-value) and distribution anomalies (per-class P95/P50 spread, per-code max/min ratio, flat-rate distributions). Exact counts are used for all validity checks (negative, zero, extreme, dimension validity). Percentile-based checks (spread, IQR) use reservoir sampling — k=5 000 global, k=1 000 per (billing_class, negotiated_type) bucket — so memory stays bounded on large files. Per-code max/min spread is exact (all rates seen, no sampling).

**Method:** Penalty-based deductions from 100.

| | |
|---|---|
| **negative_rate_pct** | `× 5 (cap 20)` |
| **zero_rate_pct** | `× 3 (cap 15)` |
| **extreme_rate_pct** | `× 5 (cap 25) — ffs only` |
| **class_spread_excess** | `(spread − threshold) × 2, max across (billing_class, negotiated_type) buckets (cap 15)` |
| **per_code_high_spread_count** | `× 0.1 (cap 15)` |
| **invalid_negotiated_type_pct** | `× 3 (cap 10) — rates silently dropped` |
| **invalid_setting_pct** | `× 1 (cap 5) — silently defaults to wildcard` |
| **invalid_severity_pct** | `× 1 (cap 5) — silently normalised to ''` |
| **institutional_with_service_codes_pct** | `× 1 (cap 5) — extra key variation` |
| **invalid_service_code_pct** | `× 2 (cap 5) — encode raises ValueError` |
| **extreme high by billing class** | professional: 25000.0, institutional: 2000000.0, both: 2000000.0, default: 500000.0 |

| | |
|---|---|
| **extreme low** | 0.01 |
| **spread warn p95 over p50 by class** | professional: 5, institutional: 10, both: 10, default: 5 |
| **per rate context max min ratio** | professional_codes: 20, facility_drg_codes: 50 |
| **flat rate iqr p75 threshold pct** | 5.0 |
| **flat rate min rates to check** | 100 |
| **spread min n to flag** | 50 |
| **per code min n to flag** | 3 |

## Dashboard: MRF Identity Key

(ingest-time — not stored in report JSON)

The dashboard assigns a persistent `mrf_key` to each MRF so that all validation runs of the same file are grouped together in the score-history view, even if the payer re-exports the file at a new URL.

| | |
|---|---|
| **Tier 1 — entity + plan_id** | Used when both `reporting_entity_name` and `plan_id` are present. Key input: `plan\|<entity>\|<plan_id_type>\|<plan_id>`. Stable across monthly re-exports. |
| **Tier 2 — URL hash** | Fallback when `plan_id` is absent. Key input: the raw file location URL/path. Entity name alone is not used — a payer publishes multiple distinct plans under the same entity name and without `plan_id` they cannot be safely distinguished. A URL change produces a different key. |

The key is a 16-character MD5 hex digest of the input string (case-insensitive, whitespace-stripped). **This run:** mrf_key = `c254098866036231` · entity = Quality Care Partners · plan_id = `311435470` (hios) · tier = 1 (entity + plan_id)

## Provider Geographic Coverage

(supplemental — does not affect score)

Geographic analysis is a supplemental feature computed on demand after scoring completes. It does not affect any scoring dimension — it is an observational overlay to assess the breadth and distribution of in-network providers.

| | |
|---|---|
| **NPPES** | CMS National Plan and Provider Enumeration System — monthly full-replacement CSV. Maps each NPI to its primary registered ZIP code. |
| **ZCTA centroids** | GeoNames US postal code file. Maps each 5-digit ZIP to a (latitude, longitude) centroid for map placement. |

**Process:** Extract all NPIs from the MRF file → resolve each NPI to its primary practice ZIP via NPPES → aggregate provider count per ZIP → map each ZIP to a lat/lon centroid via ZCTA → render as a weighted heatmap (intensity ∝ provider count per ZIP).

**Limitations:** NPIs absent from NPPES (recently issued, test NPIs, EINs) are excluded and reduce the geocoding match rate. Location reflects the provider's NPPES-registered primary address, not necessarily where they accept this specific plan. Map viewport covers the bounding box of ZIP codes representing 90% of total provider count, dropping sparse geographic outliers.