

# MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2026-03/20260218\_Signature\_Care\_EPO\_in-network\_rate.json  
Size: 4.26 GB • MD5: 26f28a054e35f5a57baf1b319ad764ad

Payer: **Signature Care** • File Date: **2026-02-18** • Generated: **2026-04-24 00:23 EDT** • Tool Version: **1.0.0** • Elapsed: **687.10s**

# 78.9

Usable w/ Caution

Errors: 0 • Warnings: 12 • Info: 4

## TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2026-03/2026-03\_plan\_ref\_f7e870d86e5a0ea993f287c2a4162c90\_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2026-03/20260218\_Signature\_Care\_EPO\_in-network\_rate.json

Plan Name	Plan ID	Issuer / Sponsor	Market
EPO	Signature Care (custom)	—	—

## CMS Official Schema Validation

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

## Dimension Scores

Dimension	Score	Weight	Findings
Schema Integrity	89.7	30%	4
Provider Mapping	96.9	15%	3
Code Coverage	70.0	15%	3
Pricing Sanity	67.3	40%	6

## Schema Integrity — Findings

---

Score: 89.7

**WARNING** header\_conditional

Partial plan fields present — missing: ['plan\_id', 'plan\_id\_type']

---

**WARNING** file\_freshness

File is 65 days old (last\_updated\_on exceeds the 45-day threshold)

---

**WARNING** expired\_prices

333568 negotiated\_prices have past expiration dates (6.5%)

---

**INFO** far\_future\_expiration

1050 expiration\_dates are more than 3 years in the future

---

## Provider Mapping — Findings

---

Score: 96.9

**WARNING** `npi_validity`

0.10% of NPIs failed Luhn checksum validation (27 of 27473)

- 0
- 1111111111
- 1111111113
- 1111111114
- 1111111115
- ... and 5 more

---

**WARNING** `ein_validity`

15 EINs failed IRS prefix validation

- 10564156
- 10570843
- 10573848
- 10672877
- 10877201
- ... and 5 more

---

**INFO** `duplicate_npis`

1977 NPIs appear in more than one provider group

---

## Code Coverage — Findings

---

Score: 70.0

**WARNING** `billing_code_format`

11304 CPT codes do not match expected format

---

**WARNING** `billing_code_format`

1550 HCPCS codes do not match expected format

---

**WARNING** `duplicate_billing_codes`

10646 billing codes appear in more than one in\_network item (99.9%)

---

## Pricing Sanity — Findings

Score: 67.3

### INFO per\_diem\_rates

167 per-diem rates (0.0%) — not dollar amounts; excluded from spread analysis

### INFO percentage\_rates

610191 percentage rates (12.0%) — values represent % of a reference rate, not dollar amounts; excluded from spread analysis

### WARNING rate\_spread\_by\_class

billing\_class='professional' / negotiated\_type='negotiated': P95/P50 spread is 8.1x (threshold: 5x, N=4,306,988 (1,000 sampled), high confidence)

### WARNING rate\_spread\_by\_class

billing\_class='institutional' / negotiated\_type='negotiated': P95/P50 spread is 51.8x (threshold: 10x, N=178,171 (1,000 sampled), high confidence)

### WARNING per\_code\_rate\_spread

678 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
J1826	HCPCS	negotiated	institutional	ffs	—	\$4.48	\$4.48	\$462.72	\$3084.40	688.5x	12
38240	CPT	negotiated	institutional	ffs	—	\$130.39	\$291.01	\$7193.77	\$34730.83	266.4x	5
J9263	HCPCS	negotiated	professional	ffs	—	\$0.05	\$0.06	\$0.17	\$12.00	240.0x	421
J9263	HCPCS	negotiated	professional	ffs	—	\$0.05	\$0.06	\$0.22	\$12.00	240.0x	229
J9263	HCPCS	negotiated	professional	ffs	—	\$0.05	\$0.06	\$0.24	\$12.00	240.0x	200
J9263	HCPCS	negotiated	professional	ffs	—	\$0.05	\$0.06	\$0.39	\$12.00	240.0x	108
J9263	HCPCS	negotiated	professional	ffs	—	\$0.05	\$0.06	\$0.28	\$11.80	236.0x	55
J9263	HCPCS	negotiated	professional	ffs	—	\$0.06	\$0.06	\$0.41	\$12.00	200.0x	131
J9263	HCPCS	negotiated	professional	ffs	—	\$0.06	\$0.06	\$0.42	\$12.00	200.0x	95
J9263	HCPCS	negotiated	professional	ffs	—	\$0.06	\$0.06	\$0.74	\$12.00	200.0x	36

### WARNING institutional\_with\_service\_codes

125830 institutional billing\_class prices (2.47%) have service\_code set. service\_code (place-of-service) is a professional billing concept; setting it on institutional rates creates extra rate-context-key variation.

## Recommended Actions

1. **provider\_mapping** `npi_validity`

P2

0.10% of NPIs failed Luhn checksum validation (27 of 27473)

2. **provider\_mapping** `ein_validity`

P2

15 EINs failed IRS prefix validation

3. **pricing** `rate_spread_by_class`

P2

billing\_class='professional' / negotiated\_type='negotiated': P95/P50 spread is 8.1x (threshold: 5x, N=4,306,988 (1,000 sampled), high confidence)

4. **pricing** `rate_spread_by_class`

P2

billing\_class='institutional' / negotiated\_type='negotiated': P95/P50 spread is 51.8x (threshold: 10x, N=178,171 (1,000 sampled), high confidence)

5. **pricing** `per_code_rate_spread`

P2

678 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.

6. **pricing** `institutional_with_service_codes`

P2

125830 institutional billing\_class prices (2.47%) have service\_code set. service\_code (place-of-service) is a professional billing concept; setting it on institutional rates creates extra rate-context-key variation.

7. **schema** `header_conditional`

P2

Partial plan fields present — missing: ['plan\_id', 'plan\_id\_type']

8. **schema** `file_freshness`

P2

File is 65 days old (last\_updated\_on exceeds the 45-day threshold)

9. **schema** expired\_prices

P2

333568 negotiated\_prices have past expiration dates (6.5%)

10. **code\_coverage** billing\_code\_format

P3

11304 CPT codes do not match expected format

11. **code\_coverage** billing\_code\_format

P3

1550 HCPCS codes do not match expected format

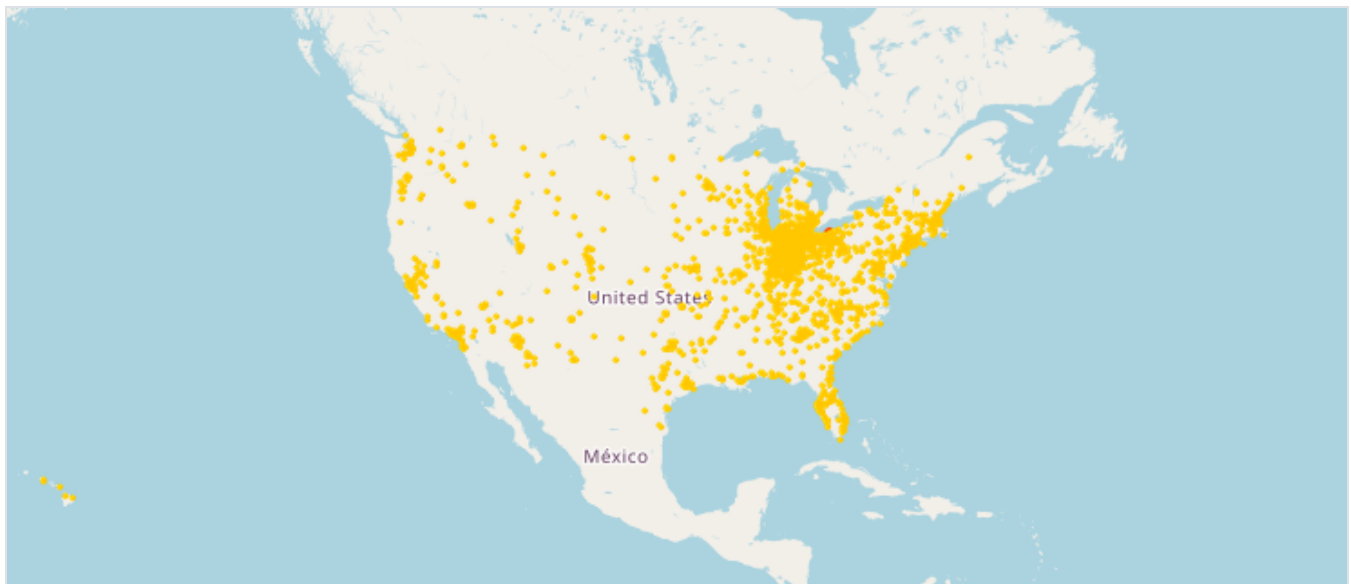
12. **code\_coverage** duplicate\_billing\_codes

P3

10646 billing codes appear in more than one in\_network item (99.9%)

## Provider Geographic Coverage

25043 unique NPIs found — 24876 geocoded (99%) — 2456 zip codes represented.



## Schema Integrity — Metrics

header\_missing\_fields

header_conditional_issues	Partial plan fields present — missing: ['plan_id', 'plan_id_type']	
file_age_days	65	
items_total	5095517	
items_missing_required_pct	0.0	
items_empty_rates	0	
prices_total	5095517	
prices_missing_required_pct	0.0	
prices_missing_field_breakdown		
prices_missing_service_code	0	
prices_invalid_billing_class	0	
rates_without_providers	0	
negotiation_arrangements	ffs	5095517
billing_code_types	CPT	3712740
	MS-DRG	5664
	RC	360
	HCPCS	1376406
	CSTM-ALL	347
expired_prices	333568	
invalid_expiration_format	0	

## Provider Mapping — Metrics

provider_references_in_file	1630
provider_group_ids_referenced	1361
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	27473
invalid_npi_count	27
invalid_npi_examples	0, 0, 0, 0, 0, 0, 0, 0, 0, 0
eins_validated	1361
invalid_ein_count	15

<b>invalid_ein_examples</b>	10564156, 10570843, 10573848, 10672877, 10877201, 10877201, 10894487, 20701635, 285485892, 30575326
<b>empty_npi_groups</b>	0
<b>groups_without_tin</b>	0
<b>npis_in_multiple_groups</b>	1977

## Code Coverage — Metrics

<b>unique_codes_total</b>	10655										
<b>duplicate_codes</b>	10646										
<b>duplicate_pct</b>	99.92										
<b>by_code_type</b>	<table border="1"> <tr> <td>CPT</td> <td>7160</td> </tr> <tr> <td>MS-DRG</td> <td>708</td> </tr> <tr> <td>RC</td> <td>49</td> </tr> <tr> <td>HCPCS</td> <td>2737</td> </tr> <tr> <td>CSTM-ALL</td> <td>1</td> </tr> </table>	CPT	7160	MS-DRG	708	RC	49	HCPCS	2737	CSTM-ALL	1
CPT	7160										
MS-DRG	708										
RC	49										
HCPCS	2737										
CSTM-ALL	1										
<b>unknown_code_types</b>											
<b>format_invalid_by_type</b>	<table border="1"> <tr> <td>CPT</td> <td>11304</td> </tr> <tr> <td>HCPCS</td> <td>1550</td> </tr> </table>	CPT	11304	HCPCS	1550						
CPT	11304										
HCPCS	1550										
<b>codes_not_in_reference</b>	reference_not_loaded										

most_frequent_codes	Type	Code	Occurrences
	CPT	90846	817
	CPT	90847	817
	CPT	90853	817
	CPT	90880	817
	CPT	90849	813
	CPT	98966	801
	CPT	98968	801
	CPT	99253	792
	CPT	99254	792
	CPT	99255	792
	CPT	90885	788
	CPT	90887	788
	CPT	99252	788
	CPT	90832	787
	CPT	90834	787
	CPT	90837	787
	CPT	90792	786
	CPT	90791	785
	CPT	90833	785
	CPT	90836	785

## Pricing Sanity — Metrics

---

total_prices_checked	5095517
total_rates	4485159
per_diem_rates	167
percentage_rates	610191
negative_rates	0
zero_rates	0
extreme_high_rates	2123
extreme_low_rates	0

<b>rate_distribution</b>	<b>sample_n</b>	4485159
	<b>sample_k</b>	5000
	<b>confidence</b>	high
	<b>p5</b>	5.0195
	<b>p25</b>	45.5
	<b>p50</b>	272.16
	<b>p75</b>	895.2624999999999
	<b>p95</b>	2537.6270000000001
	<b>p99</b>	5858.9176000000002

<b>by_billing_class</b>	<b>Class / Type</b>	<b>Count</b>	<b>Median</b>	<b>p25</b>	<b>p75</b>	<b>p95</b>	<b>Confidence</b>
	professional/ negotiated	4,306,988	267.3	46.2	839.1	2174.5	high
	institutional/ negotiated	178,171	218.4	39.7	977.2	11311.3	high

<b>negotiated_types</b>	<b>negotiated</b>	4485159
-------------------------	-------------------	---------

<b>unique_rate_contexts</b>	375861
-----------------------------	--------

<b>rate_key_dimension_validity</b>	<b>invalid_negotiated_type</b>	0
	<b>invalid_negotiated_types_seen</b>	{}
	<b>invalid_setting</b>	0
	<b>invalid_settings_seen</b>	{}
	<b>invalid_severity_of_illness</b>	0
	<b>severity_on_non_apr_drg</b>	0
	<b>institutional_with_service_codes</b>	125830
	<b>invalid_service_code_format</b>	0
	<b>billing_code_modifier_too_long</b>	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).

<b>provider_resolution (60%)</b>	$\text{resolution\_rate\%} \times 0.60$
<b>npi_validity (30%)</b>	$(100 - \text{invalid\_npi\_pct} \times 5) \times 0.30$
<b>ein_validity (10%)</b>	$10 - \min(\text{invalid\_ein\_count} \times 0.2, 10)$

## Code Coverage

Tracks every (billing\_code\_type, billing\_code) pair and flags unrecognized CMS TIC code types, format violations for CPT/HCCPS/NDC, and duplicates (same code appearing in multiple in\_network items).

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

<b>per_unknown_code_type</b>	3 pts each (cap 20)
<b>format_invalid_pct</b>	$\times 0.5$ (cap 10)
<b>duplicate_code_pct</b>	$\times 2$ (cap 20)
<b>codes_not_in_reference_pct</b>	$\times 0.5$ (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.

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

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

<b>Tier 1 — entity + plan_id</b>	Used when both <code>reporting_entity_name</code> and <code>plan_id</code> are present. Key input: <code>plan &lt;entity&gt; &lt;plan_id_type&gt; &lt;plan_id&gt;</code> . Stable across monthly re-exports.
<b>Tier 2 — URL hash</b>	Fallback when <code>plan_id</code> 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 <code>plan_id</code> 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 = ec1a18f3aafed2ca · entity = Signature Care · tier = 2 (URL hash)`

## 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.

<b>NPPES</b>	CMS National Plan and Provider Enumeration System — monthly full-replacement CSV. Maps each NPI to its primary registered ZIP code.
<b>ZCTA centroids</b>	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.