

MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2026-03/20260218_Signature_Care_ELite_in-network_rates.json
Size: 4.26 GB • MD5: 1b27a70d19804e0af3480808a0b2aeefe

Payer: **Signature Care** • File Date: **2026-02-18** • Generated: **2026-04-24 00:08 EDT** • Tool Version: **1.0.0** • Elapsed: **676.00s**

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_2ba7bf4ff6d257fd41172b638f41ff62_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2026-03/20260218_Signature_Care_ELite_in-network_rates.json

Plan Name	Plan ID	Issuer / Sponsor	Market
ELite	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

332876 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 27472)

- 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

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

INFO percentage_rates

610181 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 9.2x (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 59.7x (threshold: 10x, N=177,474 (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

125119 institutional billing_class prices (2.46%) 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 27472)

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 9.2x (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 59.7x (threshold: 10x, N=177,474 (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

125119 institutional billing_class prices (2.46%) 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

332876 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

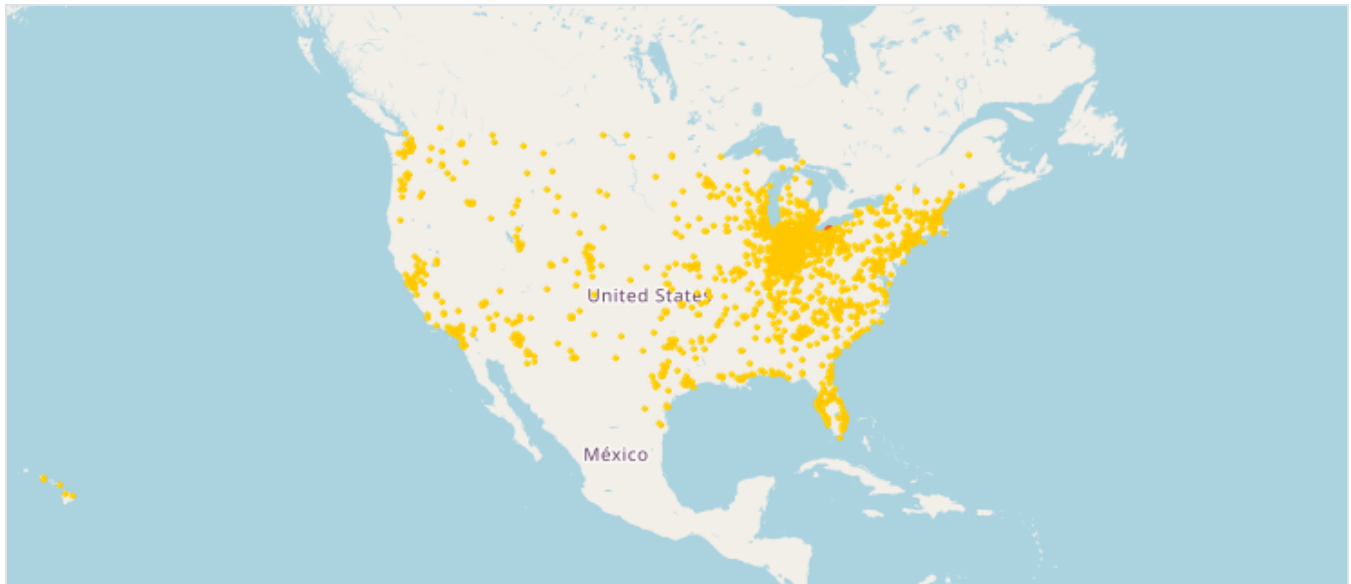
12. **code_coverage** duplicate_billing_codes

P3

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

Provider Geographic Coverage

25042 unique NPIs found — 24875 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	5094806	
items_missing_required_pct	0.0	
items_empty_rates	0	
prices_total	5094806	
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	5094806
billing_code_types	CPT	3712725
	MS-DRG	5664
	RC	353
	HCPCS	1375717
	CSTM-ALL	347
expired_prices	332876	
invalid_expiration_format	0	

Provider Mapping — Metrics

provider_references_in_file	1630
provider_group_ids_referenced	1360
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	27472
invalid_npi_count	27
invalid_npi_examples	0, 0, 0, 0, 0, 0, 0, 0, 0, 0
eins_validated	1360
invalid_ein_count	15

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

Code Coverage — Metrics

unique_codes_total	10655										
duplicate_codes	10646										
duplicate_pct	99.92										
by_code_type	<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										
unknown_code_types											
format_invalid_by_type	<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										
codes_not_in_reference	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	5094806
total_rates	4484462
per_diem_rates	163
percentage_rates	610181
negative_rates	0
zero_rates	0
extreme_high_rates	2123
extreme_low_rates	0

rate_distribution	sample_n	4484462
	sample_k	5000
	confidence	high
	p5	4.4990000000000006
	p25	47.097500000000004
	p50	260.445
	p75	866.7
	p95	2440.5465
	p99	5382.0330000000008

by_billing_class	Class / Type Count Median p25 p75 p95 Confidence						
	professional/ negotiated	4,306,988	273.4	44.8	864.3	2519.4	high
	institutional/ negotiated	177,474	200.4	29.0	990.1	11966.5	high

negotiated_types	negotiated	4484462
-------------------------	------------	---------

unique_rate_contexts	375857
-----------------------------	--------

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	125119
	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%)	$\text{resolution_rate\%} \times 0.60$
npi_validity (30%)	$(100 - \text{invalid_npi_pct} \times 5) \times 0.30$
ein_validity (10%)	$10 - \min(\text{invalid_ein_count} \times 0.2, 10)$

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.

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

negative_rate_pct	$\times 5$ (cap 20)
zero_rate_pct	$\times 3$ (cap 15)
extreme_rate_pct	$\times 5$ (cap 25) — ffs only
class_spread_excess	$(\text{spread} - \text{threshold}) \times 2$, max across (billing_class, negotiated_type) buckets (cap 15)
per_code_high_spread_count	$\times 0.1$ (cap 15)
invalid_negotiated_type_pct	$\times 3$ (cap 10) — rates silently dropped
invalid_setting_pct	$\times 1$ (cap 5) — silently defaults to wildcard
invalid_severity_pct	$\times 1$ (cap 5) — silently normalised to ''
institutional_with_service_codes_pct	$\times 1$ (cap 5) — extra key variation
invalid_service_code_pct	$\times 2$ (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 <code>reporting_entity_name</code> and <code>plan_id</code> are present. Key input: <code>plan <entity> <plan_id_type> <plan_id></code> . Stable across monthly re-exports.
Tier 2 — URL hash	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 = e011b9fb4af660a0 · 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.

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.