

MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-01-30_encore-health_encore_in-network-rates.json
Size: 431.01 MB • MD5: ad7c29dd90f2ee64389e655f76ced1e7

Payer: **The Healthcare Group dba Encore Health Network** • File Date: **2026-02-01** • Generated: **2026-04-24 12:01 EDT**
• Tool Version: **1.0.0** • Elapsed: **130.30s**

65.0

Limited Reliability

score capped at 65.0 — CMS official schema validation failed (see meta for details)

Errors: 2 • Warnings: 7 • Info: 2

TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-02_plan_ref_dbec9768b7ad2a5bedee24d3929168bf_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-01-30_encore-health_encore_in-network-rates.json

Plan Name	Plan ID	Issuer / Sponsor	Market
Encore	Encore Health (custom)	—	—

CMS Official Schema Validation

FAILED (exit code 1) — File does not conform to the CMS schema.

Validator output:

```
warn Schema version v1.0 was provided, but file indicates it conforms to schema version 1.0. v1.0 will be
used.
error Could not find a schema version named "v1.0". Available versions are:
v2.0.0
v0.1
v0.2
v0.3
v0.3.1
v0.3.2
v0.4.0
v0.4.1
v0.5.0
v0.5.3
v0.7.0
v0.8.0
v0.8.1
v0.9.0
v0.9.1
v0.10.0
v0.10.1
v0.11.0
v0.12.0
v0.13.0
v1.0.0
v0.10.3
v1.0.4
v1.0.6
v1.0.7
v1.1.0
v1.1.1
v1.1.2
v1.2.0
v1.3.0
v1.3.1
v1.3.2
v1.3.3
v1.4.0
v1.5.0
v1.6.0
v1.6.1
v1.6.2
v2.0.1
v2.1.0
v2.2.0
v2.2.1
```

Dimension Scores

Dimension	Score	Weight	Findings
Schema Integrity	70.0	30%	3
Provider Mapping	99.8	15%	3
Code Coverage	90.0	15%	1
Pricing Sanity	77.9	40%	4

Schema Integrity — Findings

Score: 70.0

WARNING `header_conditional`

plan_sponsor_name is required when plan_id_type == 'ein'

WARNING `file_freshness`

File is 82 days old (last_updated_on exceeds the 45-day threshold)

ERROR `cms_schema_validation`

CMS official schema validator FAILED (exit code 1). File does not conform to the TIC in-network-rates schema.

Provider Mapping — Findings

Score: 99.8

WARNING `npi_validity`

0.08% of NPIs failed Luhn checksum validation (42 of 53899)

- 10027007
- 100267970
- 352051598
- 364797599
- 1033190625
- ... and 5 more

WARNING `ein_validity`

0.34% of EINs failed IRS prefix validation (17 of 5004)

- 009466371
- 076587846
- 085425447
- 175426910
- 181347811
- ... and 5 more

INFO `duplicate_npis`

7018 NPIs appear in more than one provider group

Code Coverage — Findings

Score: 90.0

WARNING `billing_code_format`

6078 CPT codes do not match expected format

Pricing Sanity — Findings

Score: 77.9

INFO `percentage_rates`

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

ERROR `zero_rates`

3653 zero-dollar rates (0.44%) — CMS schema requires `negotiated_rate > 0` (`exclusiveMinimum`)

WARNING `rate_spread_by_class`

`billing_class='professional' / negotiated_type='negotiated'`: P95/P50 spread is 7.7x (threshold: 5x, N=826,862 (1,000 sampled), high confidence)

WARNING `per_code_rate_spread`

1634 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
Q2057	CPT	negotiated	professional	ffs	—	\$60.00	\$969054.65	\$1029499.65	\$2311860.00	38531.0x	12
Q2056	CPT	negotiated	professional	ffs	—	\$60.00	\$731684.90	\$819407.92	\$2303709.25	38395.2x	38
Q2042	CPT	negotiated	professional	ffs	—	\$60.00	\$757997.32	\$844458.06	\$2246229.75	37437.2x	38
Q2055	CPT	negotiated	professional	ffs	—	\$60.00	\$707077.00	\$792329.27	\$2240807.75	37346.8x	38
Q2054	CPT	negotiated	professional	ffs	—	\$60.00	\$699337.13	\$779541.72	\$2194479.75	36574.7x	38
Q2053	CPT	negotiated	professional	ffs	—	\$60.00	\$640223.45	\$724283.38	\$2083154.38	34719.2x	38
Q2041	CPT	negotiated	professional	ffs	—	\$60.00	\$673697.95	\$746028.55	\$2082699.50	34711.7x	38
J7680	CPT	negotiated	professional	ffs	—	\$0.01	\$82.50	\$70.00	\$100.00	10000.0x	6
J7681	CPT	negotiated	professional	ffs	—	\$0.01	\$82.50	\$70.00	\$100.00	10000.0x	6
P2028	CPT	negotiated	professional	ffs	—	\$0.01	\$6.93	\$10.13	\$85.00	8500.0x	29

Recommended Actions

1. **schema** `cms_schema_validation`

P1

CMS official schema validator FAILED (exit code 1). File does not conform to the TIC in-network-rates schema.

2. **pricing** zero_rates

P1

3653 zero-dollar rates (0.44%) — CMS schema requires negotiated_rate > 0 (exclusiveMinimum)

3. **provider_mapping** npi_validity

P2

0.08% of NPIs failed Luhn checksum validation (42 of 53899)

4. **provider_mapping** ein_validity

P2

0.34% of EINs failed IRS prefix validation (17 of 5004)

5. **pricing** rate_spread_by_class

P2

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 7.7x (threshold: 5x, N=826,862 (1,000 sampled), high confidence)

6. **pricing** per_code_rate_spread

P2

1634 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.

7. **schema** header_conditional

P2

plan_sponsor_name is required when plan_id_type == 'ein'

8. **schema** file_freshness

P2

File is 82 days old (last_updated_on exceeds the 45-day threshold)

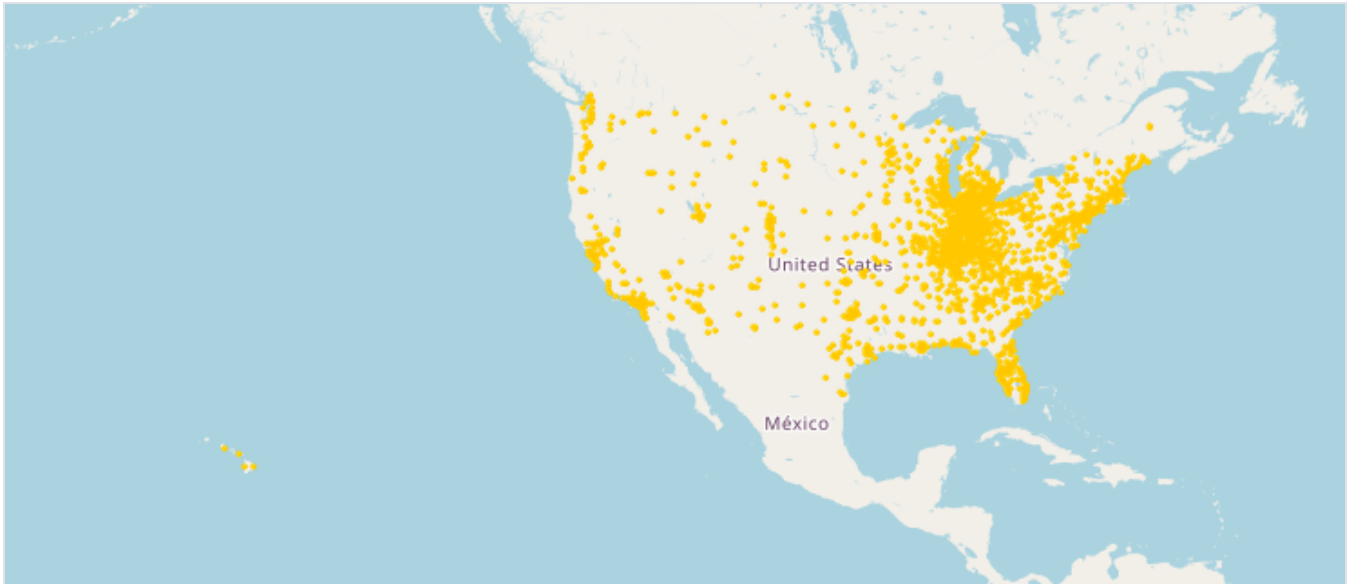
9. **code_coverage** billing_code_format

P3

6078 CPT codes do not match expected format

Provider Geographic Coverage

43967 unique NPIs found — 43588 geocoded (99%) — 3345 zip codes represented.



Schema Integrity — Metrics

header_missing_fields		
header_conditional_issues	plan_sponsor_name is required when plan_id_type == 'ein'	
file_age_days	82	
items_total	16247	
items_missing_required_pct	0.0	
items_empty_rates	0	
prices_total	830782	
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	16247

billing_code_types	CPT	15476
	AP-DRG	770
	CSTM-ALL	1
expired_prices	0	
invalid_expiration_format	0	

Provider Mapping — Metrics

provider_references_in_file	5191
provider_group_ids_referenced	5004
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	53899
invalid_npi_count	42
npi_validity_rate_pct	99.92
invalid_npi_examples	10027007, 100267970, 352051598, 364797599, 1033190625, 1033190625, 1033190625, 1033190625, 1036073393, 1111111111
eins_validated	5004
invalid_ein_count	17
ein_validity_rate_pct	99.66
invalid_ein_examples	009466371, 076587846, 085425447, 175426910, 181347811, 184366929, 189663275, 196446983, 284500476, 287447787
empty_npi_groups	0
groups_without_tin	0
npis_in_multiple_groups	7018

Code Coverage — Metrics

unique_codes_total	16247
duplicate_codes	0
duplicate_pct	0.0

by_code_type	CPT	15476	
	AP-DRG	770	
	CSTM-ALL	1	
unknown_code_types			
format_invalid_by_type	CPT	6078	
codes_not_in_reference	reference_not_loaded		
most_frequent_codes	Type	Code	Occurrences
	CPT	0001U	1
	CPT	0002M	1
	CPT	0002U	1
	CPT	0003M	1
	CPT	0003U	1
	CPT	0004M	1
	CPT	0005U	1
	CPT	0006M	1
	CPT	0007M	1
	CPT	0007U	1
	CPT	0008U	1
	CPT	0009U	1
	CPT	0010U	1
	CPT	0011M	1
	CPT	0011U	1
	CPT	0012M	1
	CPT	0013M	1
	CPT	0015M	1
	CPT	0016M	1
CPT	0016U	1	

Pricing Sanity — Metrics

total_prices_checked	830782
total_rates	830712
per_diem_rates	0
percentage_rates	70
negative_rates	0
zero_rates	3653

extreme_high_rates	638																					
extreme_low_rates	0																					
rate_distribution	<table border="1"> <tr> <td>sample_n</td> <td>830712</td> </tr> <tr> <td>sample_k</td> <td>5000</td> </tr> <tr> <td>confidence</td> <td>high</td> </tr> <tr> <td>p5</td> <td>13.639000000000001</td> </tr> <tr> <td>p25</td> <td>80.0</td> </tr> <tr> <td>p50</td> <td>321.95500000000004</td> </tr> <tr> <td>p75</td> <td>983.6575</td> </tr> <tr> <td>p95</td> <td>2943.751500000001</td> </tr> <tr> <td>p99</td> <td>7103.671500000004</td> </tr> </table>	sample_n	830712	sample_k	5000	confidence	high	p5	13.639000000000001	p25	80.0	p50	321.95500000000004	p75	983.6575	p95	2943.751500000001	p99	7103.671500000004			
sample_n	830712																					
sample_k	5000																					
confidence	high																					
p5	13.639000000000001																					
p25	80.0																					
p50	321.95500000000004																					
p75	983.6575																					
p95	2943.751500000001																					
p99	7103.671500000004																					
by_billing_class	<table border="1"> <thead> <tr> <th>Class / Type</th> <th>Count</th> <th>Median</th> <th>p25</th> <th>p75</th> <th>p95</th> <th>Confidence</th> </tr> </thead> <tbody> <tr> <td>professional/ negotiated</td> <td>826,862</td> <td>350.1</td> <td>95.8</td> <td>1041.3</td> <td>2690.9</td> <td>high</td> </tr> <tr> <td>institutional/ negotiated</td> <td>3,850</td> <td>30875.6</td> <td>20084.9</td> <td>53052.6</td> <td>125447.2</td> <td>high</td> </tr> </tbody> </table>	Class / Type	Count	Median	p25	p75	p95	Confidence	professional/ negotiated	826,862	350.1	95.8	1041.3	2690.9	high	institutional/ negotiated	3,850	30875.6	20084.9	53052.6	125447.2	high
Class / Type	Count	Median	p25	p75	p95	Confidence																
professional/ negotiated	826,862	350.1	95.8	1041.3	2690.9	high																
institutional/ negotiated	3,850	30875.6	20084.9	53052.6	125447.2	high																
negotiated_types	<table border="1"> <tr> <td>negotiated</td> <td>830712</td> </tr> </table>	negotiated	830712																			
negotiated	830712																					
unique_rate_contexts	126803																					
rate_key_dimension_validity	<table border="1"> <tr> <td>invalid_negotiated_type</td> <td>0</td> </tr> <tr> <td>invalid_negotiated_types_seen</td> <td>{}</td> </tr> <tr> <td>invalid_setting</td> <td>0</td> </tr> <tr> <td>invalid_settings_seen</td> <td>{}</td> </tr> <tr> <td>invalid_severity_of_illness</td> <td>0</td> </tr> <tr> <td>severity_on_non_apr_drg</td> <td>0</td> </tr> <tr> <td>institutional_with_service_codes</td> <td>0</td> </tr> <tr> <td>invalid_service_code_format</td> <td>0</td> </tr> <tr> <td>billing_code_modifier_too_long</td> <td>0</td> </tr> </table>	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			
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%)	$\text{resolution_rate\%} \times 0.60$
npi_validity (30%)	$(100 - \text{invalid_npi_pct} \times 5) \times 0.30$
ein_validity (10%)	$10 - (\text{invalid_ein_pct} \times 0.1)$ [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/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 = 55ee3ec3f939ada1 · entity = The Healthcare Group dba Encore Health Network · plan_id = 352067676 (EIN) · 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 \propto 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.