MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-02-01_midlands-choice_2_in-network-rates.json.gz

Size: 36.93 MB • MD5: 37b9c7c7976f01894d6807141ca9877f

Payer: **Midlands Choice** • File Date: **2026-02-05** • Generated: **2026-04-24 12:08 EDT** • Tool Version: **1.0.0** • Elapsed: **311.50s**

65.0

Limited Reliability

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

Errors: 2 • Warnings: 11 • Info: 2

TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-02_plan_ref_e7cd5f5c82c1c63f41b07fb716400106_index.json • Midlands Choice (Third Party)

Canonical: s3://talon-storage-private/mrf-feed-uploads/2026-02/2026-02-01_midlands-choice_2_in-network-rates.json.gz

Plan Name	Plan ID	Issuer / Sponsor	Market
MidlandsChoice2	2 (EIN)	—	group

CMS Official Schema Validation

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

Validator output:

```
warn  Schema version v2.1 was provided, but file indicates it conforms to schema version 2.1. v2.1 will be
used.
error Could not find a schema version named "v2.1". 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	100.0	15%	2
Code Coverage	70.0	15%	3
Pricing Sanity	59.8	40%	7

Schema Integrity — Findings

Score: 70.0

WARNING `header_conditional`

plan_sponsor_name is required when plan_id_type == 'ein'

WARNING `file_freshness`

File is 78 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: 100.0

WARNING `ein_validity`

0.01% of EINs failed IRS prefix validation (1 of 11681)

- 293253974

INFO `duplicate_npis`

19140 NPIs appear in more than one provider group

Code Coverage — Findings

Score: 70.0

WARNING `billing_code_format`

67492 CPT codes do not match expected format

WARNING `billing_code_format`

556 HCPCS codes do not match expected format

WARNING `duplicate_billing_codes`

24251 billing codes appear in more than one in_network item (99.0%)

Pricing Sanity — Findings

Score: 59.8

INFO `percentage_rates`

432847 percentage rates (28.6%) — values represent % of a reference rate, not dollar amounts; excluded from spread analysis

ERROR `zero_rates`

20362 zero-dollar rates (1.88%) — CMS schema requires negotiated_rate > 0 (exclusiveMinimum)

WARNING `rate_spread_by_class`

billing_class='both' / negotiated_type='negotiated': P95/P50 spread is 11.9x (threshold: 10x, N=344,164 (1,000 sampled), high confidence)

WARNING `rate_spread_by_class`

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 12.1x (threshold: 5x, N=534,870 (1,000 sampled), high confidence)

WARNING `rate_spread_by_class`

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 12.8x (threshold: 10x, N=203,175 (1,000 sampled), high confidence)

WARNING `per_code_rate_spread`

1220 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
J9052	HCPCS	negotiated	institutional	ffs	both	$0.07	$0.12	$55.81	$493.43	**7049.0×**	22
J9052	HCPCS	negotiated	institutional	ffs	both	$0.07	$0.12	$55.81	$493.43	**7049.0×**	22
J9052	HCPCS	negotiated	institutional	ffs	both	$0.07	$0.12	$55.81	$493.43	**7049.0×**	22
A9582	HCPCS	negotiated	professional	ffs	both	$1.44	$3448.60	$3614.08	$6100.85	**4236.7×**	30
A9582	HCPCS	negotiated	professional	ffs	both	$1.44	$3448.60	$3614.08	$6100.85	**4236.7×**	30
A9582	HCPCS	negotiated	professional	ffs	both	$1.44	$3448.60	$3614.08	$6100.85	**4236.7×**	30
A4210	HCPCS	negotiated	professional	ffs	both	$1.21	$941.90	$938.63	$1571.90	**1299.1×**	30
A4210	HCPCS	negotiated	professional	ffs	both	$1.21	$941.90	$938.63	$1571.90	**1299.1×**	30
A4210	HCPCS	negotiated	professional	ffs	both	$1.21	$941.90	$938.63	$1571.90	**1299.1×**	30
A9541	HCPCS	negotiated	professional	ffs	both	$1.02	$143.89	$304.09	$1068.58	**1047.6×**	16

WARNING `institutional_with_service_codes`

216767 institutional billing_class prices (14.31%) 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. **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

20362 zero-dollar rates (1.88%) — CMS schema requires negotiated_rate > 0 (exclusiveMinimum)

3. **provider_mapping** `ein_validity` — P2

0.01% of EINs failed IRS prefix validation (1 of 11681)

4. **pricing** `rate_spread_by_class` — P2

billing_class='both' / negotiated_type='negotiated': P95/P50 spread is 11.9x (threshold: 10x, N=344,164 (1,000 sampled), high confidence)

5. **pricing** `rate_spread_by_class` — P2

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 12.1x (threshold: 5x, N=534,870 (1,000 sampled), high confidence)

6. **pricing** `rate_spread_by_class` — P2

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 12.8x (threshold: 10x, N=203,175 (1,000 sampled), high confidence)

7. **pricing** `per_code_rate_spread` — P2

1220 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. **pricing** `institutional_with_service_codes` **P2**

216767 institutional billing_class prices (14.31%) 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.

9. **schema** `header_conditional` **P2**

plan_sponsor_name is required when plan_id_type == 'ein'

10. **schema** `file_freshness` **P2**

File is 78 days old (last_updated_on exceeds the 45-day threshold)

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

67492 CPT codes do not match expected format

12. **code_coverage** `billing_code_format` **P3**

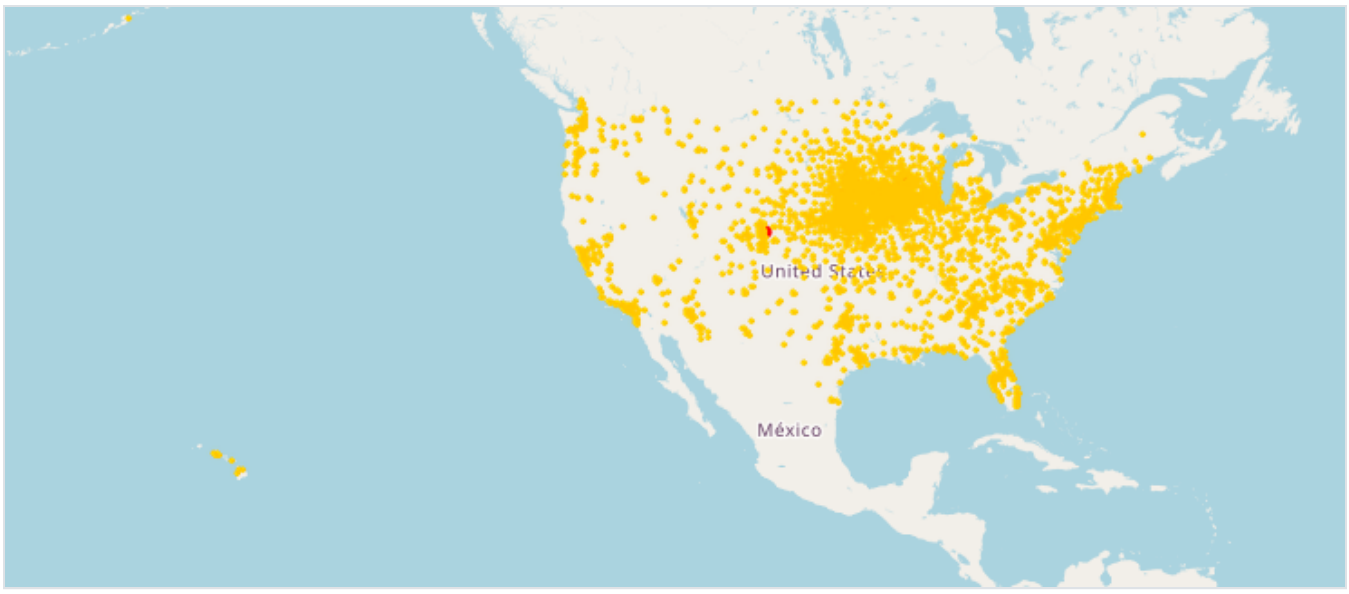
556 HCPCS codes do not match expected format

13. **code_coverage** `duplicate_billing_codes` **P3**

24251 billing codes appear in more than one in_network item (99.0%)

Provider Geographic Coverage

55335 unique NPIs found — 55305 geocoded (100%) — 4073 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	78	
items_total	1150562	
items_missing_required_pct	0.0	
items_empty_rates	0	
prices_total	1515056	
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**	1150562
billing_code_types	**HCPCS**	435388
	CPT	711757
	MS-DRG	3243
	RC	174
expired_prices	0	

invalid_expiration_format	0

Provider Mapping — Metrics

provider_references_in_file	12047
provider_group_ids_referenced	11961
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	98702
invalid_npi_count	0
npi_validity_rate_pct	100.0
invalid_npi_examples	
eins_validated	11681
invalid_ein_count	1
ein_validity_rate_pct	99.99
invalid_ein_examples	293253974
empty_npi_groups	0
groups_without_tin	0
npis_in_multiple_groups	19140

Code Coverage — Metrics

unique_codes_total	24501	
duplicate_codes	24251	
duplicate_pct	98.98	
by_code_type	HCPCS	10710
	CPT	12890
	MS-DRG	803
	RC	98
unknown_code_types		
format_invalid_by_type	CPT	67492
	HCPCS	556

codes_not_in_reference	reference_not_loaded

most_frequent_codes

Type	Code	Occurrences
CPT	78216	86
CPT	78215	86
CPT	78202	86
CPT	78201	86
CPT	78199	86
CPT	78195	86
CPT	78191	86
CPT	78185	86
CPT	78140	86
CPT	78130	86
CPT	78122	86
CPT	78121	86
CPT	78120	86
CPT	78111	86
CPT	72197	86
CPT	72196	86
CPT	78264	86
CPT	78262	86
CPT	78261	86
CPT	78258	86

Pricing Sanity — Metrics

total_prices_checked	1515056
total_rates	1082209
per_diem_rates	0
percentage_rates	432847
negative_rates	0
zero_rates	20362
extreme_high_rates	673
extreme_low_rates	0

rate_distribution

sample_n	1082209
sample_k	5000
confidence	high
p5	5.3895
p25	80.7225
p50	368.255
p75	1404.115
p95	4857.032500000001
p99	13316.32500000003

by_billing_class

Class / Type	Count	Median	p25	p75	p95	Confidence
both/ negotiated	344,164	365.9	67.8	1149.0	4337.1	high
professional/ negotiated	534,870	306.7	72.9	1040.6	3713.5	high
institutional/ negotiated	203,175	646.0	81.0	2605.6	8261.0	high

negotiated_types

negotiated	1082209

unique_rate_contexts 298156

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	216767
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 = `4c0ad2f73b129863` · entity = Midlands Choice · plan_id = `47-0804331` (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 ∝ 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.