MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2026-01/2026-01-01_imagine360-Imagine_in-network-rates_fixed.json

Size: 26.60 MB • MD5: 1c62ed682bb1e49d3a8e1dc0ccb9f02c

Payer: **Imagine360** • File Date: **2025-12-19T18:48:29** • Generated: **2026-04-24 14:12 EDT** • Tool Version: **1.0.0** • Elapsed: **29.40s**

74.8

Limited Reliability

Errors: 1 • Warnings: 10 • Info: 1

TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2026-01/2026-01_plan_ref_1614f2a35f85d240aced30ab96373088_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2026-01/2026-01-01_imagine360-Imagine_in-network-rates_fixed.json

Plan Name	Plan ID	Issuer / Sponsor	Market
Elap-imagine360	Performance Health (custom)	—	—
Imagine Health Network	—	—	group

CMS Official Schema Validation

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

Dimension Scores

Dimension	Score	Weight	Findings
Schema Integrity	61.0	30%	4
Provider Mapping	99.8	15%	3
Code Coverage	98.5	15%	1
Pricing Sanity	66.9	40%	4

Schema Integrity — Findings

Score: 61.0

WARNING `header_conditional`

last_updated_on '2025-12-19T18:48:29' is not a valid YYYY-MM-DD date

WARNING `header_conditional`

Partial plan fields present — missing: ['plan_id', 'plan_id_type']

ERROR `item_required_fields`

100.00% of in_network items are missing required fields

WARNING `expired_prices`

119563 negotiated_prices have past expiration dates (100.0%)

Provider Mapping — Findings

Score: 99.8

WARNING `npi_validity`

0.16% of NPIs failed Luhn checksum validation (141 of 89147)

- 0
- 179003553
- 1011354773
- 1013123018
- 1013123019
- *… and 5 more*

WARNING `ein_validity`

0.06% of EINs failed IRS prefix validation (3 of 4661)

- 078566668
- 088343029
- 491761946

INFO `duplicate_npis`

6506 NPIs appear in more than one provider group

Code Coverage — Findings

Score: 98.5

WARNING `billing_code_format`

180 HCPCS codes do not match expected format

Pricing Sanity — Findings

Score: 66.9

WARNING `high_frequency_rate_value`

1 rate value(s) appear with suspiciously high frequency (≥0.5% of dollar rates and ≥50 occurrences) — likely placeholder/sentinel values rather than real negotiated rates.

- {'rate': 0.01, 'count': 3709, 'pct': 3.1}

WARNING `rate_spread_by_class`

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 8.5x (threshold: 5x, N=89,688 (1,000 sampled), high confidence)

WARNING `rate_spread_by_class`

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 21.3x (threshold: 10x, N=29,875 (1,000 sampled), high confidence)

WARNING `per_code_rate_spread`

783 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
S2900	HCPCS	negotiated	institutional	ffs	—	$0.01	$0.22	$1964.20	$8666.02	**866602.0×**	15
S2900	HCPCS	negotiated	professional	ffs	—	$0.01	$825.00	$1503.11	$4500.00	**450000.0×**	13
J3490	HCPCS	negotiated	professional	ffs	—	$0.01	$5.54	$92.86	$3113.17	**311317.0×**	40
J7307	HCPCS	negotiated	professional	ffs	—	$0.01	$1335.76	$1253.53	$2550.73	**255073.0×**	35
J3590	HCPCS	negotiated	professional	ffs	—	$0.01	$0.01	$256.59	$1500.00	**150000.0×**	6
90670	CPT	negotiated	professional	ffs	—	$0.01	$257.81	$257.79	$1126.00	**112600.0×**	121
90710	CPT	negotiated	professional	ffs	—	$0.01	$299.36	$272.85	$645.23	**64523.0×**	92
90651	CPT	negotiated	professional	ffs	—	$0.01	$301.90	$301.74	$624.64	**62464.0×**	171
90620	CPT	negotiated	professional	ffs	—	$0.01	$250.00	$231.96	$497.70	**49770.0×**	61
G0121	HCPCS	negotiated	professional	ffs	—	$0.01	$309.78	$253.12	$449.56	**44956.0×**	3

Recommended Actions

1. **schema** `item_required_fields` **P1**

100.00% of in_network items are missing required fields

2. provider_mapping npi_validity P2

0.16% of NPIs failed Luhn checksum validation (141 of 89147)

3. provider_mapping ein_validity P2

0.06% of EINs failed IRS prefix validation (3 of 4661)

4. pricing high_frequency_rate_value P2

1 rate value(s) appear with suspiciously high frequency (≥0.5% of dollar rates and ≥50 occurrences) — likely placeholder/sentinel values rather than real negotiated rates.

5. pricing rate_spread_by_class P2

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 8.5x (threshold: 5x, N=89,688 (1,000 sampled), high confidence)

6. pricing rate_spread_by_class P2

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 21.3x (threshold: 10x, N=29,875 (1,000 sampled), high confidence)

7. pricing per_code_rate_spread P2

783 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 '2025-12-19T18:48:29' is not a valid YYYY-MM-DD date

9. schema header_conditional P2

Partial plan fields present — missing: ['plan_id', 'plan_id_type']

10. **schema** `expired_prices` **P2**

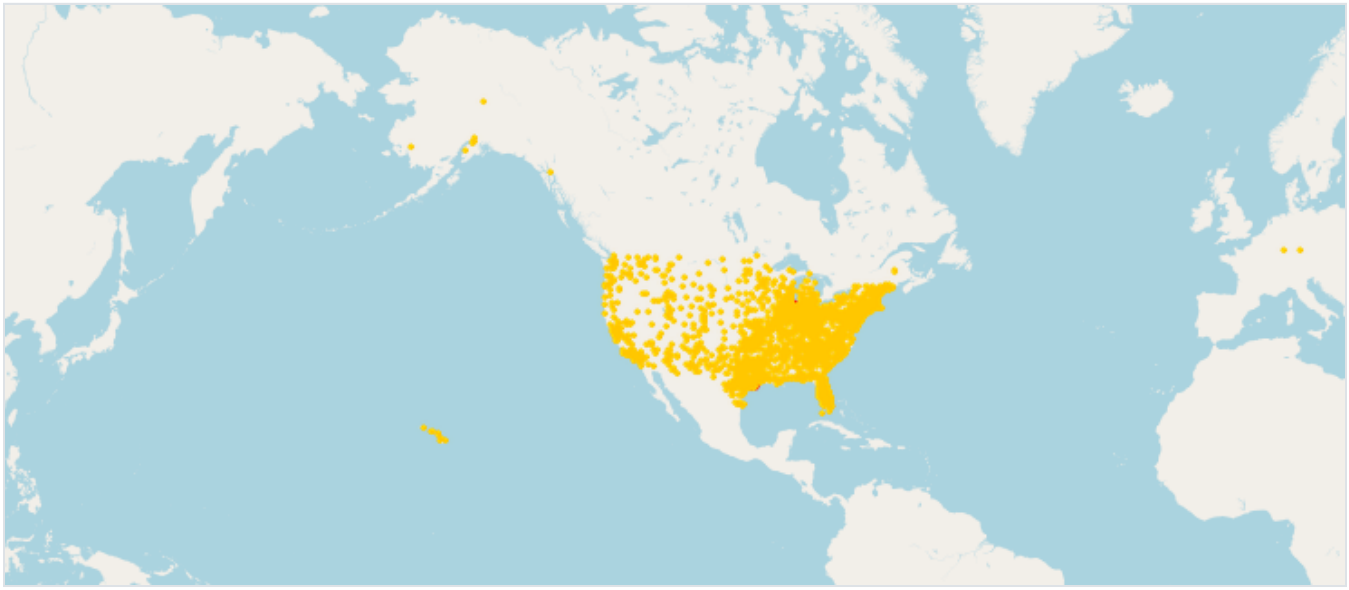
119563 negotiated_prices have past expiration dates (100.0%)

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

180 HCPCS codes do not match expected format

Provider Geographic Coverage

81877 unique NPIs found — 81494 geocoded (100%) — 5769 zip codes represented.



Schema Integrity — Metrics

header_missing_fields	
header_conditional_issues	last_updated_on '2025-12-19T18:48:29' is not a valid YYYY-MM-DD date, Partial plan fields present — missing: ['plan_id', 'plan_id_type']
file_age_days	None
items_total	5944
items_missing_required_pct	100.0
items_empty_rates	0
prices_total	119563

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	5944
billing_code_types	HCPCS	1566
	CPT	4166
	RC	212
expired_prices	119563	
invalid_expiration_format	0	

Provider Mapping — Metrics

provider_references_in_file	8289
provider_group_ids_referenced	4661
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	89147
invalid_npi_count	141
npi_validity_rate_pct	99.84
invalid_npi_examples	0, 179003553, 1011354773, 1013123018, 1013123019, 1013123020, 1013123021, 1013123022, 1013286890, 1058835828
eins_validated	4661
invalid_ein_count	3
ein_validity_rate_pct	99.94
invalid_ein_examples	078566668, 088343029, 491761946
empty_npi_groups	0
groups_without_tin	0
npis_in_multiple_groups	6506

Code Coverage — Metrics

unique_codes_total	5944	
duplicate_codes	0	
duplicate_pct	0.0	
by_code_type	HCPCS	1566
	CPT	4166
	RC	212
unknown_code_types		
format_invalid_by_type	HCPCS	180
codes_not_in_reference	reference_not_loaded	

most_frequent_codes

Type	Code	Occurrences
HCPCS	M1007	1
HCPCS	A4206	1
HCPCS	G9247	1
HCPCS	A5057	1
HCPCS	A5056	1
HCPCS	A4435	1
CPT	99492	1
CPT	99494	1
HCPCS	A4213	1
HCPCS	E2361	1
HCPCS	C8924	1
HCPCS	G9368	1
HCPCS	A4208	1
HCPCS	G9226	1
CPT	77295	1
CPT	76376	1
CPT	76377	1
CPT	90649	1
HCPCS	S5010	1
HCPCS	J7121	1

Pricing Sanity — Metrics

total_prices_checked	119563

total_rates	119563
per_diem_rates	0
percentage_rates	0
negative_rates	0
zero_rates	0
extreme_high_rates	0
extreme_low_rates	0

rate_distribution

sample_n	119563
sample_k	5000
confidence	high
p5	0.9295000000000002
p25	22.345
p50	91.86
p75	206.07999999999998
p95	1512.7965000000006
p99	5983.50930000001

by_billing_class

Class / Type	Count	Median	p25	p75	p95	Confidence
`professional/negotiated`	89,688	73.6	16.7	162.1	629.1	high
`institutional/negotiated`	29,875	225.5	58.3	1043.8	4814.3	high

negotiated_types

negotiated	119563

unique_rate_contexts	15540

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 = `c8bf01c2aa2cb2cf` · entity = Imagine360 · 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.