

MRF Quality Report

s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12-01_imagine360-Imagine_in-network-rates_fixed.json
Size: 26.37 MB • MD5: dbc668a69b9e92c465ff03c9c2f5a65

Payer: **Imagine360** • File Date: **2025-11-25T20:55:33** • Generated: **2026-04-24 20:23 EDT** • Tool Version: **1.0.0** • Elapsed: **18.80s**

75.3

Usable w/ Caution

Errors: 1 • Warnings: 10 • Info: 1

TOC Plan References

TOC: s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12_plan_ref_1614f2a35f85d240aced30ab96373088_index.json • Talon

Canonical: s3://talon-storage-private/mrf-feed-uploads/2025-12/2025-12-01_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	68.3	40%	4

Schema Integrity — Findings

Score: 61.0

WARNING header_conditional

last_updated_on '2025-11-25T20:55:33' 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

118441 negotiated_prices have past expiration dates (100.0%)

Provider Mapping — Findings

Score: 99.8

WARNING `npi_validity`

0.15% of NPIs failed Luhn checksum validation (137 of 88879)

- 0
- 179003553
- 1011354773
- 1013123018
- 1013123019
- ... and 5 more

WARNING `ein_validity`

0.06% of EINs failed IRS prefix validation (3 of 4639)

- 078566668
- 088343029
- 491761946

INFO `duplicate_npis`

6551 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: 68.3

WARNING `high_frequency_rate_value`

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

- `{'rate': 0.01, 'count': 3716, 'pct': 3.14}`

WARNING `rate_spread_by_class`

billing_class='professional' / negotiated_type='negotiated': P95/P50 spread is 8.6x (threshold: 5x, N=88,901 (1,000 sampled), high confidence)

WARNING `rate_spread_by_class`

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 16.8x (threshold: 10x, N=29,540 (1,000 sampled), high confidence)

WARNING `per_code_rate_spread`

768 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	866602.0x	15
S2900	HCPCS	negotiated	professional	ffs	—	\$0.01	\$825.00	\$1502.05	\$4500.00	450000.0x	13
J3490	HCPCS	negotiated	professional	ffs	—	\$0.01	\$5.54	\$92.88	\$3113.17	311317.0x	40
J7307	HCPCS	negotiated	professional	ffs	—	\$0.01	\$1335.76	\$1253.02	\$2550.73	255073.0x	35
J3590	HCPCS	negotiated	professional	ffs	—	\$0.01	\$0.01	\$256.59	\$1500.00	150000.0x	6
90670	CPT	negotiated	professional	ffs	—	\$0.01	\$260.67	\$257.95	\$1126.00	112600.0x	121
90710	CPT	negotiated	professional	ffs	—	\$0.01	\$299.42	\$273.62	\$645.23	64523.0x	93
90651	CPT	negotiated	professional	ffs	—	\$0.01	\$301.90	\$301.48	\$624.64	62464.0x	170
90620	CPT	negotiated	professional	ffs	—	\$0.01	\$248.38	\$231.15	\$497.70	49770.0x	60
G0121	HCPCS	negotiated	professional	ffs	—	\$0.01	\$309.78	\$253.12	\$449.56	44956.0x	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

0.15% of NPIs failed Luhn checksum validation (137 of 88879)

3. **provider_mapping** `ein_validity`

P2

0.06% of EINs failed IRS prefix validation (3 of 4639)

4. **pricing** `high_frequency_rate_value`

P2

1 rate value(s) appear with suspiciously high frequency ($\geq 0.5\%$ 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 8.6x (threshold: 5x, N=88,901 (1,000 sampled), high confidence)

6. **pricing** `rate_spread_by_class`

P2

billing_class='institutional' / negotiated_type='negotiated': P95/P50 spread is 16.8x (threshold: 10x, N=29,540 (1,000 sampled), high confidence)

7. **pricing** `per_code_rate_spread`

P2

768 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 '2025-11-25T20:55:33' 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

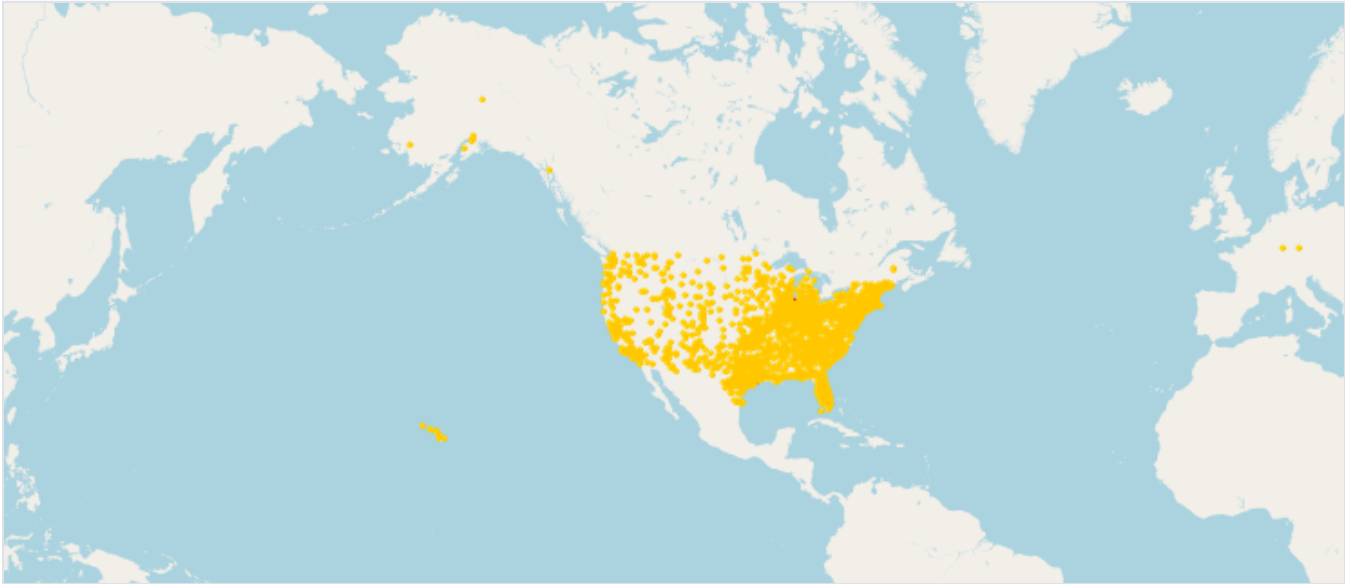
118441 negotiated_prices have past expiration dates (100.0%)

11. **code_coverage** billing_code_format P3

180 HCPCS codes do not match expected format

Provider Geographic Coverage

81566 unique NPIs found — 81183 geocoded (100%) — 5764 zip codes represented.



Schema Integrity — Metrics

header_missing_fields	
header_conditional_issues	last_updated_on '2025-11-25T20:55:33' is not a valid YYYY-MM-DD date, Partial plan fields present — missing: ['plan_id', 'plan_id_type']
file_age_days	None
items_total	5913
items_missing_required_pct	100.0
items_empty_rates	0
prices_total	118441

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	5913
billing_code_types	HCPCS	1552
	CPT	4148
	RC	213
expired_prices	118441	
invalid_expiration_format	0	

Provider Mapping — Metrics

provider_references_in_file	8250
provider_group_ids_referenced	4639
unresolved_references	0
resolution_rate_pct	100.0
npis_validated	88879
invalid_npi_count	137
npi_validity_rate_pct	99.85
invalid_npi_examples	0, 179003553, 1011354773, 1013123018, 1013123019, 1013123020, 1013123021, 1013123022, 1013286890, 1058835828
eins_validated	4639
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	6551

Code Coverage — Metrics

unique_codes_total	5913		
duplicate_codes	0		
duplicate_pct	0.0		
by_code_type	HCPCS	1552	
	CPT	4148	
	RC	213	
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	118441
----------------------	--------

total_rates	118441																					
per_diem_rates	0																					
percentage_rates	0																					
negative_rates	0																					
zero_rates	0																					
extreme_high_rates	0																					
extreme_low_rates	0																					
rate_distribution	<table border="1"> <tr> <td>sample_n</td> <td>118441</td> </tr> <tr> <td>sample_k</td> <td>5000</td> </tr> <tr> <td>confidence</td> <td>high</td> </tr> <tr> <td>p5</td> <td>2.0285</td> </tr> <tr> <td>p25</td> <td>23.34</td> </tr> <tr> <td>p50</td> <td>93.59</td> </tr> <tr> <td>p75</td> <td>209.4425</td> </tr> <tr> <td>p95</td> <td>1461.5700000000002</td> </tr> <tr> <td>p99</td> <td>6465.65040000000095</td> </tr> </table>	sample_n	118441	sample_k	5000	confidence	high	p5	2.0285	p25	23.34	p50	93.59	p75	209.4425	p95	1461.5700000000002	p99	6465.65040000000095			
sample_n	118441																					
sample_k	5000																					
confidence	high																					
p5	2.0285																					
p25	23.34																					
p50	93.59																					
p75	209.4425																					
p95	1461.5700000000002																					
p99	6465.65040000000095																					
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>88,901</td> <td>68.4</td> <td>17.2</td> <td>151.2</td> <td>589.5</td> <td>high</td> </tr> <tr> <td>institutional/ negotiated</td> <td>29,540</td> <td>221.5</td> <td>68.6</td> <td>893.8</td> <td>3720.3</td> <td>high</td> </tr> </tbody> </table>	Class / Type	Count	Median	p25	p75	p95	Confidence	professional/ negotiated	88,901	68.4	17.2	151.2	589.5	high	institutional/ negotiated	29,540	221.5	68.6	893.8	3720.3	high
Class / Type	Count	Median	p25	p75	p95	Confidence																
professional/ negotiated	88,901	68.4	17.2	151.2	589.5	high																
institutional/ negotiated	29,540	221.5	68.6	893.8	3720.3	high																
negotiated_types	<table border="1"> <tr> <td>negotiated</td> <td>118441</td> </tr> </table>	negotiated	118441																			
negotiated	118441																					
unique_rate_contexts	15401																					
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 = d60b63e79cd2411f · 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.