MEDICOM | RESEARCH

Fewer Than 1 in 10 Institutions Describe the Same Exam the Same Way

A network-scale measurement of procedure-naming standardization across 124 US healthcare organizations.

MEDICOM RESEARCH · JULY 2026

ABSTRACT

Background. Procedure descriptions underpin prior-study retrieval, hanging protocols, duplicate detection, and AI in radiology, yet their consistency is rarely measured where it matters most: across institutional boundaries.

Methods. We analyzed 22,143 prior imaging searches over a recent seven-day period across 124 US provider organizations on the Medicom network, yielding 404,748 procedure-description instances and 17,262 distinct strings. DICOM Modality metadata from 1,345,763 objects received through Medicom's SMART on FHIR applications, a largely distinct customer population, provided a second vantage point.

Results. Six routine procedures appeared under 1,258 distinct descriptions. Only 9.1 percent of organization pairs shared even one exact description for the same concept. Among descriptions naming a sided body part, 28.2 percent omitted laterality; contrast phase appeared in ten notations; the Modality field carried 121 distinct values.

Conclusions. Procedure vocabulary remains deeply local, and internal catalogs understate the problem because it becomes visible only at the network boundary. Systems that consume descriptions should match on clinical meaning, and standardization programs should measure drift where exchange happens.

INTRODUCTION

Radiology has spent two decades building shared vocabularies for how imaging procedures should be named. Logical Observation Identifiers Names and Codes (LOINC) and the Radiological Society of North America (RSNA) RadLex Playbook define a common language, DICOM constrains the metadata that travels with every study, and nearly every health system maintains a curated internal procedure catalog. What nobody has been able to measure is how well those efforts hold up at the moment a patient's imaging history crosses from one institution to another.

Medicom operates at that crossing point, moving imaging among hospitals, imaging centers, and specialty practices every day, and that traffic offers a view of standardization no single institution can produce from inside its own walls. We asked a simple question of that traffic: when institutions exchange imaging, how consistent is the language that travels with it?

DATA AND METHODS

This analysis draws on two Medicom vantage points serving largely distinct customer populations with partial overlap. The first is Medicom Smart Search, our AI-powered prior-study retrieval application. Over a recent seven-day period, users at 30 sampled organizations ran 22,143 prior imaging searches against 113 partner institutions. Of the 30 sampled organizations, 19 sit on both sides of that exchange and 11 only initiate; 94 additional organizations only respond, for 124 distinct organizations in all. Each search carries the description of the study and returns descriptions of priors held elsewhere, a mean of 17.28 per search;[1] returned and launching descriptions together yield 404,748 instances containing 17,262 distinct strings.

The second vantage point is the flow of studies into Medicom's cloud platform through our SMART on FHIR applications: 1,345,763 DICOM objects sampled from those arriving from provider organizations, satellite locations, patient uploads, and imported media. Concept-level analyses group description strings by pattern matching; agreement between organizations is measured as the share of organization pairs whose description sets intersect on at least one exact string. Full definitions appear in Methodology and notes, and network composition appears in Appendix A. No patient identifiers appear anywhere in this analysis.

[1] The median is 8 priors per search. The mean exceeds the median because a subset of searches — for patients with extensive imaging histories and complex disease — return disproportionately many priors, pulling the average upward. This subgroup, and what it suggests about which patients benefit most from Smart Search, is a natural direction for future analysis.

RESULTS

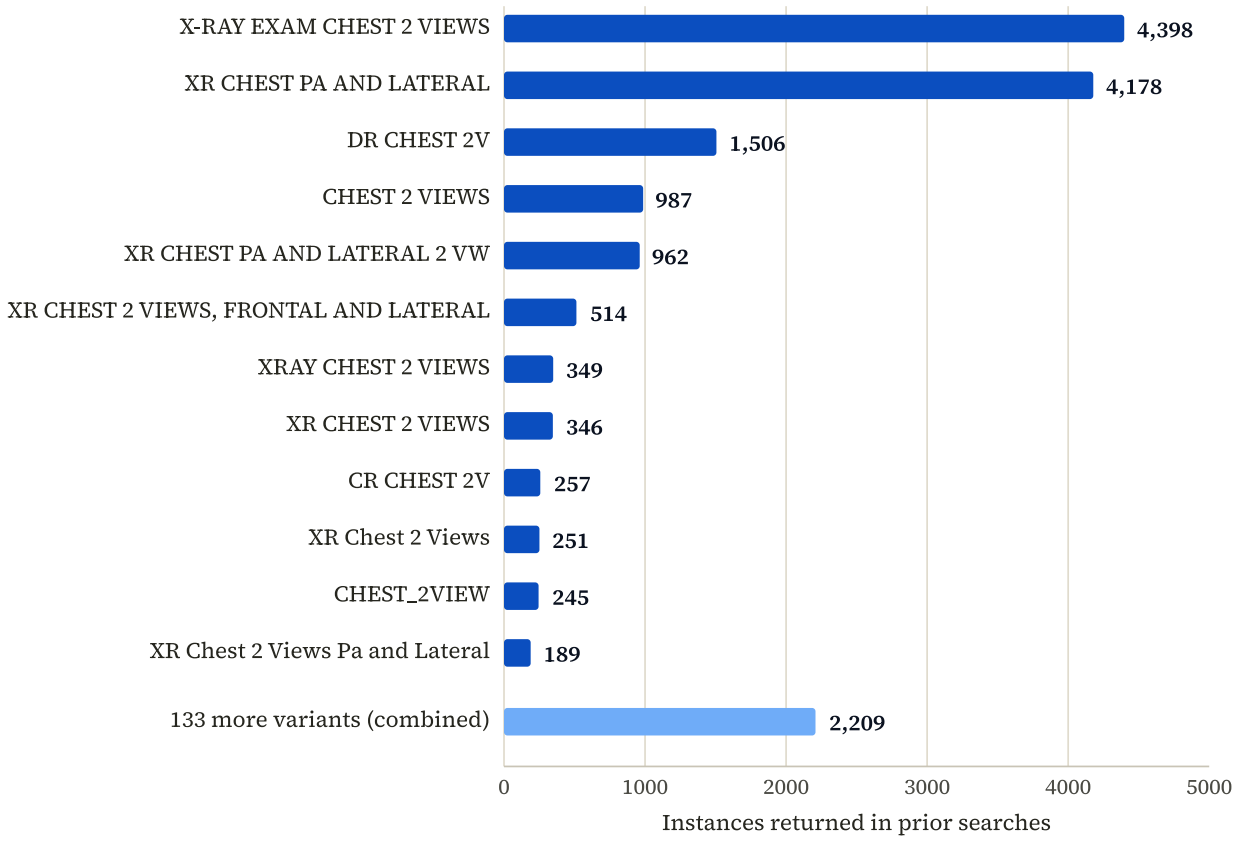
Vocabulary Proliferation

The two-view chest radiograph, among the most common exams in medicine, appeared under **145 distinct spellings** across 78 organizations (Figure 1). The distribution is heavily long-tailed: the most common spelling covers 26.8 percent of instances, the top three cover 61.5 percent, and eight cover 80 percent, yet complete coverage requires all 145 (Figure 2). The median organization uses two spellings for this exam internally (interquartile range 1 to 4); one imaging network uses 19.

Every routine procedure we examined behaves the same way: screening mammography appeared under 392 distinct names across 70 organizations, CT of the abdomen and pelvis with contrast under 234, MRI of the lumbar spine without contrast under 204, CT of the head without contrast under 148, and pelvic ultrasound under 135, totaling **1,258 distinct strings** for six routine procedures (Figure 3).

One exam, 145 spellings: the two-view chest radiograph

16,391 instances across 78 organizations. Exact strings as they travel on the network.

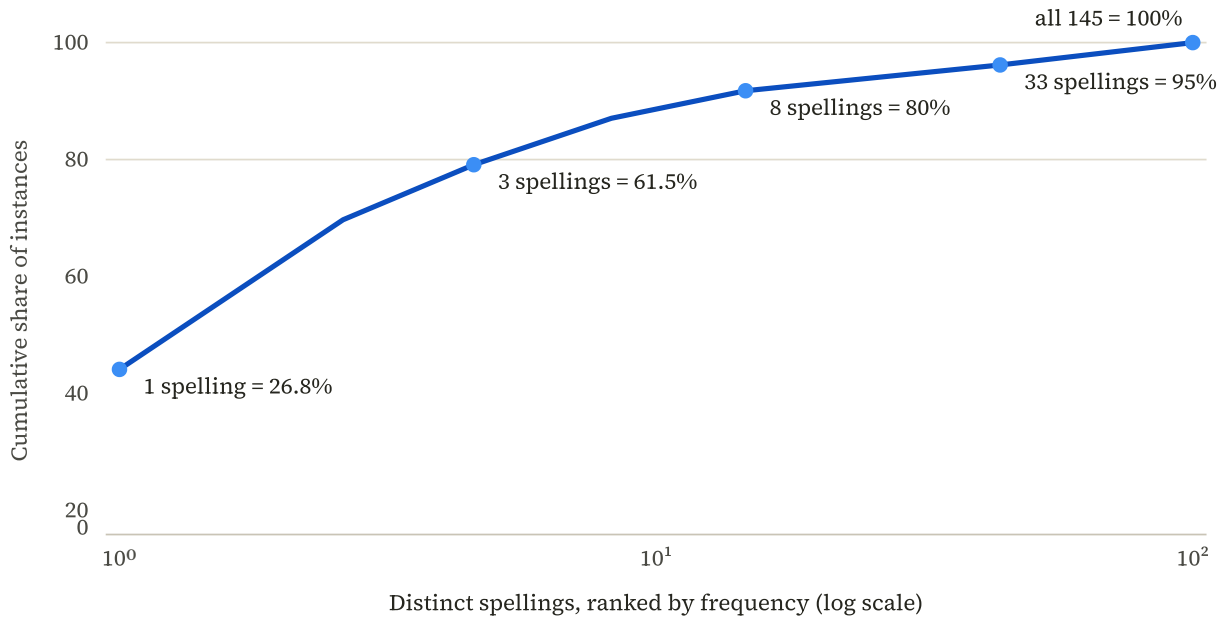


The most common spelling covers 26.8% of volume. No spelling reaches one third of the network. The 133 rarest spellings still carry 2,209 instances.

Figure 1. The 12 most frequent spellings of a two-view chest radiograph, with 133 additional variants combined. No spelling reaches one third of network volume. Source: Medicom Smart Search, seven-day sample, 2026. 404,748 procedure-description instances across 124 connected organizations.

Eight spellings cover 80 percent; full coverage takes all 145

Cumulative coverage of two-view chest radiograph instances by ranked spelling.

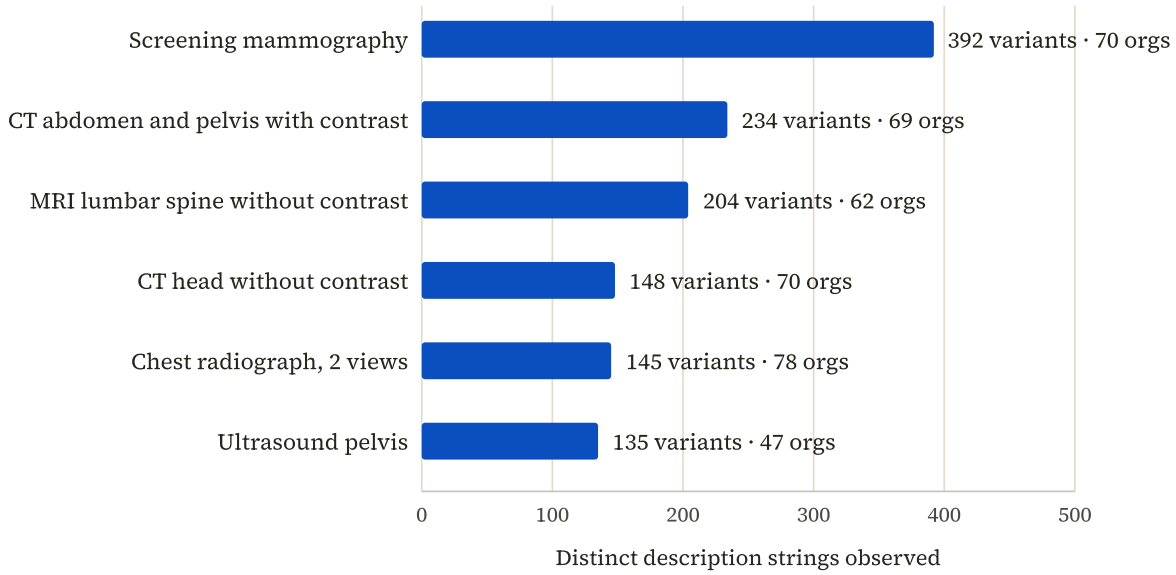


Any system that matches on exact strings faces this curve: the last 20% of volume hides behind 137 additional spellings.

Figure 2. Cumulative coverage of chest-radiograph instances by ranked spelling. Exact-string systems face a long tail: the final 20 percent of volume requires 137 additional spellings. Source: Medicom Smart Search, seven-day sample, 2026. 404,748 procedure-description instances across 124 connected organizations.

Six routine procedures, 1,258 different names

Distinct exact strings per procedure concept in one week of network traffic.



Screening mammography is the most fragmented: 392 names, an average of 5.6 per organization that performs it.

Figure 3. Distinct description strings observed per procedure concept in one week of network activity. Source: Medicom Smart Search, seven-day sample, 2026. 404,748 procedure-description instances across 124 connected organizations.

Agreement Between Institutions Is the Exception

This may be the study's central finding. Among 13,151 pairs of organizations holding descriptions for the same concept, only **9.1 percent** share even one identical string, ranging from 3.0 percent for screening mammography to 15.2 percent for lumbar-spine MRI (Figure 4). A randomly chosen pair of institutions performing the same routine exam will, more than nine times in ten, describe it with no spelling in common.

Fewer than 1 in 10 institution pairs share a single exact spelling

13,151 pairwise comparisons among organizations holding descriptions for the same concept.

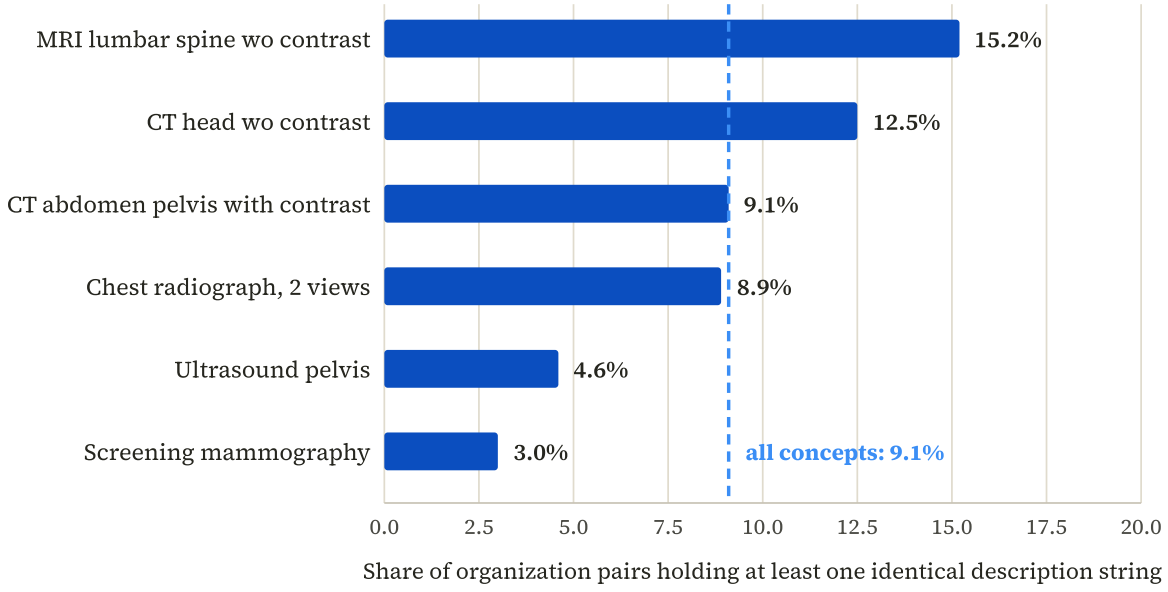


Figure 4. Share of organization pairs sharing at least one exact description string, by concept. Source: Medicom Smart Search, seven-day sample, 2026. 404,748 procedure-description instances across 124 connected organizations.

Clinically Material Attributes

The variation reaches attributes with direct clinical weight. Contrast phase appears in at least ten notations across the 96,545 instances carrying one: WO, W/O, WITHOUT, and W O mean without contrast (73,666 instances), while W WO, WITH AND WITHOUT, W/WO, W AND WO, WWO, and W&WO mean with and without (22,879 instances; Figure 5). Laterality fares worse: among 60,256 instances naming a sided body part such as a wrist, knee, or shoulder, 17,013, or 28.2 percent, carry no side marker at all (Figure 6).

Ten ways to write the contrast phase

A string match tuned to any single notation misses most of the network.

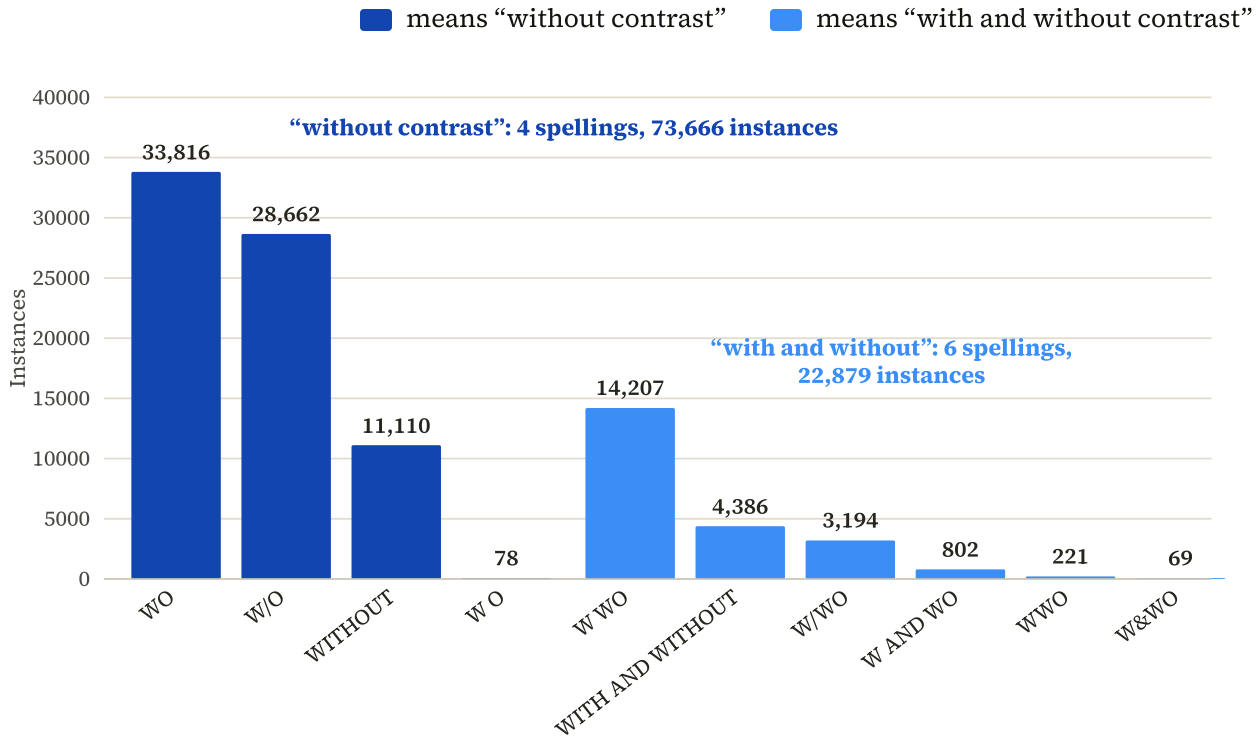
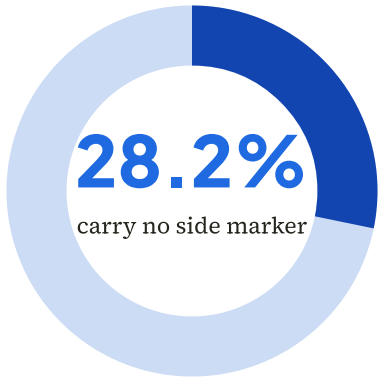


Figure 5. Ten notations for the contrast phase. A string match tuned to any single notation misses most of the network. Source: Medicom Smart Search, seven-day sample, 2026. 404,748 procedure-description instances across 124 connected organizations.

Sided body part, missing side

60,256 instances naming a lateralizable musculoskeletal body part.



More than 1 in 4 sided exams travel without a stated side. Those that do state one split across RT/LT, RIGHT/LEFT, and BILAT/BILATERAL styles.

Laterality absent (17,013) **Laterality stated (43,243)**

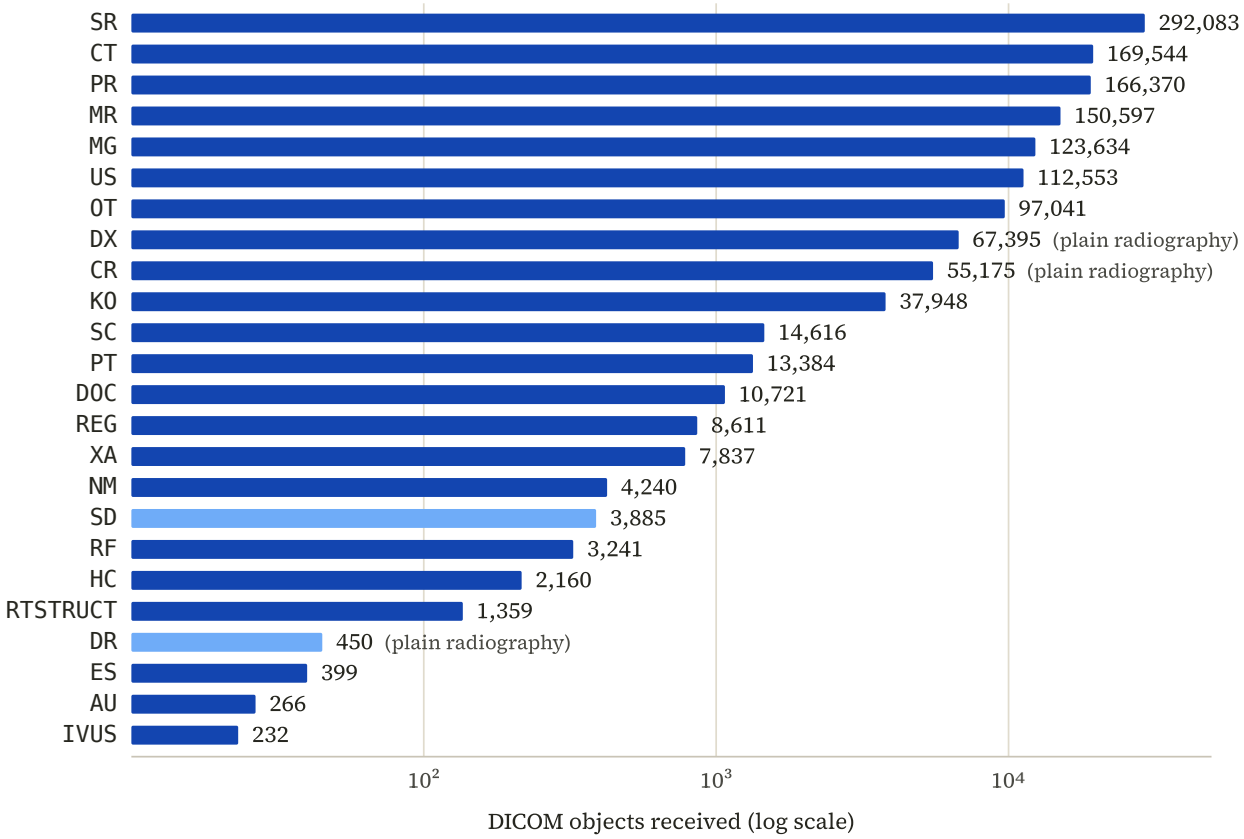
Figure 6. Share of description instances naming a sided musculoskeletal body part that state a side. Source: Medicom Smart Search, seven-day sample, 2026. Sided body parts: wrist, shoulder, knee, hip, ankle, elbow, hand, foot, and similar. Breast imaging excluded.

Drift in the Most Constrained Field

Modality is supposed to draw from a defined list of a few dozen DICOM terms. Across the 1,345,763 DICOM objects sampled from those received through our SMART on FHIR applications, we observed **121 distinct Modality values**, including lowercase variants, spelled-out phrases, vendor strings, SUPPLIES, UNKNOWN, and raw device identifiers (Figure 7). Truly invalid values are rare by volume, well under one percent of objects in every source channel; the finding is the sheer variety, and the fragmentation of single exam types across multiple valid codes. Plain radiography alone arrives under three: CR (55,175 objects), DX (67,395), and DR (450). The same feed shows that 46.8 percent of received objects are structured reports, presentation states, key object selections, and other evidence objects, so description text increasingly labels far more than pixels. If the most tightly governed attribute in the standard drifts this far in the wild, free-text descriptions never stood a chance.

121 distinct modality codes in a field built for a few dozen

Top 24 values shown. One exam type, plain radiography, fragments across CR, DX, and DR: 123,020 objects under three codes.



Long tail includes lowercase mr, mri, dx, vendor strings, SUPPLIES, UNKNOWN, REQUEST, and raw device identifiers. Light blue bars sit outside the DICOM defined term list. Invalid values are rare by volume, under 1% in every source channel; the finding is variety and fragmentation. 46.8% of received objects are reports, annotations, and other evidence objects rather than image series.

Figure 7. The 24 most frequent of 121 distinct DICOM Modality values received; flagged values sit outside the DICOM defined term list. Source: DICOM objects received via Medicom SMART on FHIR applications. 1,345,763 objects, 121 distinct Modality values.

Institutions Disagree With Themselves

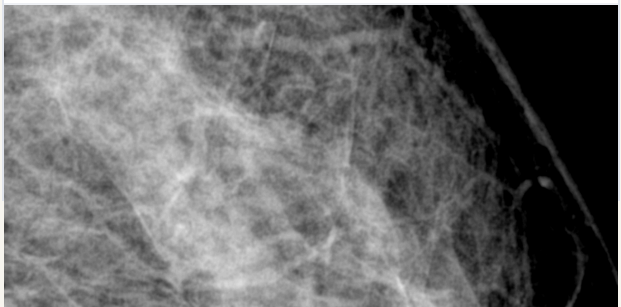
We found **962 cases** in which a single organization writes the same normalized procedure more than one way, such as one academic health system holding both *MRI CERVICAL SPINE WO CONTRAST* and *MRI SPINE CERVICAL W/O CONTRAST* in active use. The problem also resists trivial cleanup: after folding case, punctuation, spacing, and common abbreviations, only 16.4 percent of distinct strings collapse into another. The remaining **83.6 percent** of variation reflects substantive differences in local vocabulary that no formatting script can reconcile.

DISCUSSION

Every one of these strings eventually lands in front of a person or a system that must decide what it means. A radiologist reading a screening mammogram needs the right priors on the screen. A hanging protocol keys on the description to arrange them. A duplicate ordering check compares a new order against history. An AI model trains and makes inferences directly on the description string. Each of these breaks quietly when the vocabulary shifts at an institutional boundary, and the patients most affected are exactly the ones whose care crosses those boundaries: trauma transfers, oncology referrals, second opinions, and anyone whose imaging lives in more than one place.

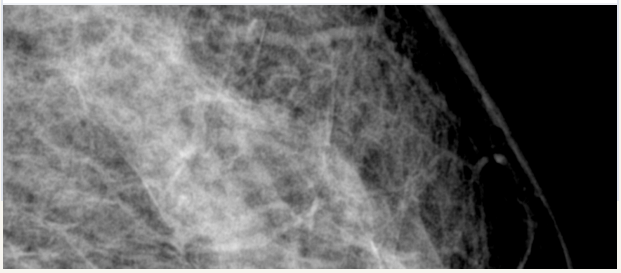
The pairwise-agreement result sharpens the point. With 9.1 percent of institution pairs sharing even one exact string for the same exam, exact matching between any two institutions fails as the default rather than the edge case. This is why relevance inside Medicom Smart Search is computed on clinical meaning: in production it returns a mean of 17.3 candidate priors per search and judges a mean of 3.6 relevant, a 79 percent reduction. A search launched from one organization's `BI MAMMO SCREEN W TOMO` phrasing surfaces a partner's `BI MAMMOGRAM DIAGNOSTIC DIGITAL TOMOSYNTHESIS BILATERAL`, because matching on spelling would leave 392 dialects of a screening mammogram invisible to one another.

HOSPITAL A · ON FILE



≠

HOSPITAL B · SEARCHES FOR



BI MAMMO SCREEN W TOMO

NO EXACT MATCH

BI MAMMOGRAM DIAGNOSTIC DIGITAL TOMOSYNTHESIS BILATERAL

The disorder is also invisible from inside any single institution. Internal catalogs look tidy, agreement within an organization is high, and the fragmentation only materializes when studies cross organizational lines. That asymmetry explains why two decades of standardization work can coexist with the numbers above: the drift accumulates precisely where no one institution can see or govern it.

RECOMMENDATIONS FOR THE INDUSTRY

Five practices follow directly from these measurements.

01 Map at the Boundary:

Standard terminologies such as the LOINC RSNA Radiology Playbook deliver the most value when applied where studies cross organizational lines. Exchange infrastructure should attach a mapped standard code to every study it moves, while preserving the local description for provenance.

02 Make Laterality and Contrast Structured:

Both attributes have dedicated homes in DICOM and in order-entry workflows. Requiring them as discrete, validated fields at ordering and at the modality worklist closes the largest clinical gaps we measured: 28.2 percent of sided-exam descriptions omit a side, and contrast phase travels in ten notations. Laterality is harder to enforce than it first appears: most systems lack the logic to determine which procedures are sided in the first place, so standardization has to precede validation — first establishing which standard procedures require a laterality value, then requiring it, ideally at ordering rather than only at the modality worklist, where errors are costlier to unwind. CMS could reinforce this requirement through billing rules, giving health systems a regulatory as well as clinical reason to close the gap.

03 Measure Vocabulary Drift as a Quality Metric:

Health systems already track report turnaround and radiation dose; distinct descriptions per procedure concept deserve the same discipline, tracked internally and against what arrives from exchange partners. Doing so requires health systems to adopt ontologies such as RadLex and present interoperability platforms with the mapping from their internal compendium to standardized codes. Internal inconsistency is material: we found 962 cases of one organization spelling the same procedure multiple ways.

04 Treat Relevancy as Infrastructure:

Prior retrieval, hanging protocols, duplicate detection, and AI cohort selection all consume descriptions, and each should match on clinical meaning. Teams procuring enterprise imaging platforms should test retrieval across vocabularies directly, because fewer than 1 in 10 institution pairs share a single exact spelling for the same exam.

05 Preserve Source Fidelity:

Normalize toward lower entropy, but annotate rather than overwrite. Reducing the variety of ways a single exam concept is described is the goal — every merge of two dialects into one shared term is a genuine reduction in entropy. But the reduction should happen by annotation, not by overwriting: a receiving system needs both the local dialect and the shared translation to remain auditable, and improving the mappings themselves requires the original text.

LIMITATIONS

Instance-level percentages are search-weighted: descriptions attached to frequently searched patients count each time they appear, so figures reflect what retrieval encounters in practice and should be read in that frame. The search observation window is seven days, and single-week sampling may underrepresent rare procedures. Concept clusters are defined by pattern matching on free text and may include or exclude edge variants. The two data populations overlap only partially, and neither is a representative sample of United States imaging as a whole. Finally, exact string comparison understates semantic agreement: 16.4 percent of distinct strings collapse under simple normalization, though the remaining 83.6 percent of variation is substantive.

METHODOLOGY AND NOTES

Search data come from Medicom Smart Search query logs covering a recent seven-day period in 2026: 22,143 prior imaging search events from 30 requesting organizations against 113 searched organizations, with 19 organizations appearing on both sides (124 distinct in all). Parsing the returned prior-study lists, a mean of 17.3 descriptions per search, yields 404,748 procedure-description instances and 17,262 distinct strings.

Pairwise agreement is the share of organization pairs, among organizations holding at least one description for a concept, whose description sets intersect on at least one exact string (13,151 pairs across six concepts). Coverage curves rank distinct strings by frequency and accumulate instance share. Relevance counts parse the priors returned per search and those judged relevant by the Smart Search relevance engine (means of 17.3 and 3.6).

Object data cover 1,345,763 DICOM objects received into Medicom's cloud platform through its SMART on FHIR applications between October 2024 and July 2026, arriving from provider organizations, satellite locations, patient uploads, and imported media, and spanning 121 distinct values of the DICOM Modality attribute. The Smart Search and SMART on FHIR application customer bases overlap only partially; the two datasets describe related but distinct networks, and neither is a representative sample of United States imaging as a whole.

Laterality analysis covers descriptions naming a sided musculoskeletal body part; breast imaging is excluded because bilateral acquisition is often implied. Normalization folds case, punctuation, spacing, and common abbreviation synonyms. No patient-level data appears in this analysis.

APPENDIX A · NETWORK REFERENCE

The 124 organizations in this sample — 30 initiating organizations and the partner institutions they searched against — are summarized below by category. Categories are deliberately broad so that no participant can be identified; no organization is named by identity, city, or state.

ORGANIZATION CATEGORY	COUNT
Community, regional & rural hospitals	57
Integrated delivery networks & multiregion systems	22
Outpatient imaging providers	17
Academic health systems & teaching hospitals	12
Oncology & other specialty groups	8
Orthopedic specialty groups	6
Federal health facilities	2
Total	**124**

Network roles: 11 organizations only initiate searches, 94 only respond, and 19 do both. Participants span more than 15 states across all four US census regions and range from quaternary academic medical centers to critical-access hospitals. Activity is concentrated: The five most active initiators account for 59 percent of search volume, and the most active single health system accounts for 29 percent.