IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

MONDO Disease Mappings: Equivalence Rules and Release Review

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-11-28 · mondo disease mappings · mondo disease ontology · mondo equivalence rules · skos exactmatch · sssom mappings · mondo omim orphanet mapping · mondo release notes · obsolete mondo terms · disease ontology crosswalk · biotech data integration

Original article: https://intuitionlabs.ai/articles/mondo-disease-mappings-equivalence-rules-release-review



In this article

1. Executive Summary
2. Introduction and Background
3. Predicate and Equivalence Rules: What Each Mapping Relation Permits
4. Key Changes: Release Review of September to November 2026
5. Source Vocabularies: OMIM, Orphanet, MeSH and ICD
6. Choosing and Freezing a Release Artifact
7. Implementation Considerations and Process Changes
8. Data Analysis and Evidence
9. Frequently Asked Questions (FAQs)
10. Implications and Future Directions
11. Conclusion

Executive Summary

Mondo (the Monarch Disease Ontology) is a monthly-released, curated disease ontology that merges disease concepts from OMIM, Orphanet, the Disease Ontology and other resources into one identifier space (mondo.monarchinitiative.org) (escholarship.org). For a data steward joining disease-labelled evidence across resources, the central question is not whether a Mondo identifier exists for a source term, but which relation sits behind the match. Mondo's own FAQ states that MONDO:equivalentTo and skos:exactMatch "conceptually overlap but they have entirely different semantics", and that equivalentTo is much stronger (mondo.monarchinitiative.org). A database cross-reference, an exact synonym, a SKOS (Simple Knowledge Organization System) match and an OWL (Web Ontology Language) equivalence are four different objects. A mapping contract for research software should treat only the strongest of them as license to pool evidence automatically.

This report sets out a predicate-by-predicate reuse matrix (Table 1), a release-change review of the three most recent Mondo releases (Table 2), and a ledger of worked cases drawn from the published release notes (Table 3). As of 28 November 2026 the latest release is v2026-11-03 (api.github.com). Its notes list 39 new terms, 265 renamings and 36 obsoletions with replacement (github.com). The release before it, v2026-10-06, listed 40 new terms, 27 renamings, 13 obsoletions with replacement and 2 without (github.com). These are published release-change counts. They are not counts of affected records in any reader's dataset.

Three rules follow from the evidence. First, pin the release tag and the artifact, because the OBO (Open Biological and Biomedical Ontologies) Foundry versioning principle requires that official release files not change after publication (obofoundry.org). Second, a replacement annotation on an obsoleted term is a pointer for review, not a proof of biological equivalence: Mondo itself separates "term replaced by" from "consider" and does not allow both on one term (mondo.readthedocs.io). Third, keep disease entities apart from phenotypic features, because the Human Phenotype Ontology (HPO) describes phenotypic abnormalities and Mondo excludes terms that represent only observable characteristics (obophenotype.github.io) (mondo.readthedocs.io).

Professional reuse is documented but should not be over-read. ClinGen's educational page states that MONDO identifiers are required in its curation interfaces, although its video dates from 2017 (clinicalgenome.org). NCBI announced in 2020 that ClinVar and the Genetic Testing Registry accept MONDO identifiers for conditions (ncbiinsights.ncbi.nlm.nih.gov). The output of this analysis is a mapping contract for research software. It is not clinical coding, diagnosis or trial-eligibility advice, and a crosswalk never supplies diagnostic equivalence.

Introduction and Background

Biotech data teams rarely work with one disease vocabulary. A target-validation table may carry OMIM numbers, a rare-disease registry may use ORPHAcodes, a literature index may use MeSH headings, and a billing-adjacent source may use ICD-10 codes. Each vocabulary draws its boundaries differently. Mondo was built to reconcile such boundaries: its 2022 preprint reported that it harmonized knowledge from 17 disease resources (www.medrxiv.org), and the 2025 Genetics paper described it as a logic-based ontology reconciling 21 disease resources (escholarship.org). The Mondo documentation describes "curated precise 1:1 equivalence axioms connecting to other resources" as the aim of the logical layer (mondo.readthedocs.io).

The word "mapping" hides at least five different assertions. This report separates them, because the choice between "join automatically", "keep as a broader or narrower association" and "send to curation" depends on which assertion is on the table. It then applies that separation to concrete release changes. The editorial asset is a mapping-policy ledger and a release-specific table of changes. It does not replace the ontology, and it does not replace the official release notes.

The audience is rare-disease and oncology data stewards, translational research leads, indication-landscape analysts and scientific platform engineers. The decision they face is whether a proposed cross-resource match is suitable for evidence aggregation, must remain a broader or narrower association, or needs curation before it supports an indication-level research decision. Evidence comes from official Mondo documentation, the project's GitHub release record, standards bodies, peer-reviewed papers and the vocabulary owners' pages. Counts quoted from live pages were read on 28 November 2026.

The process recommendation in this article, that disease-mapping decisions be written down as versioned, reviewable records rather than embedded in join code, is the author's editorial view.

Predicate and Equivalence Rules: What Each Mapping Relation Permits

Five different objects behind the word "mapping"

The first object is the **database cross-reference** (xref). The SSSOM (Simple Standard for Sharing Ontology Mappings) paper warns that xrefs "are semantically ambiguous and are used in vastly different ways in different ontologies" (pmc.ncbi.nlm.nih.gov), and that although they often correspond to exact matches, "they can correspond to broad, narrow, close or related matches" (pmc.ncbi.nlm.nih.gov). A bare xref therefore cannot be used as an evidence-join key.

The second object is the **synonym**. A synonym is a label attached to one Mondo class; it is not a relation to a class in another resource. Text matching alone is a known hazard: the 2022 preprint cautions that matching on text "can be misleading" (www.medrxiv.org).

The third object is the **SKOS mapping predicate**. Under SKOS, skos:exactMatch is the only mapping property declared transitive (www.w3.org), all other SKOS mapping properties are not transitive (www.w3.org), and exactMatch, closeMatch and relatedMatch are symmetric (www.w3.org). The SSSOM predicate guidance describes exactMatch as meaning the two terms "are intended to refer to the same thing" (mapping-commons.github.io), and says of closeMatch that it is a hazy category "and should be avoided in practice" (mapping-commons.github.io).

The fourth object is **logical equivalence**. In the Mondo editors guide, MONDO:equivalentTo "is interpreted strongly as an OWL equivalence axiom" (mondo.readthedocs.io). The FAQ goes further: Mondo uses skos:exactMatch "to bridge semantic spaces which conceptually overlap" (mondo.monarchinitiative.org), while applying strong OWL semantics with owl:equivalentClass to such loose overlaps "is, usually, inappropriate" (mondo.monarchinitiative.org). Under OWL 2, two classes are equivalent if they contain exactly the same individuals (www.w3.org). The SSSOM guidance adds that when mappings are loaded into an OWL reasoner, "even a bit of noise can have huge consequences for reasoning" (mapping-commons.github.io).

The fifth object is the **hierarchy relation** inside a resource. It is a relation among classes of one ontology and says nothing about which classes in another resource are the same. No source reviewed here licenses treating a parent class as an alias of its child.

Predicate-by-predicate reuse matrix

Table 1 below converts these definitions into a reuse contract. "Allowed reuse" is the most permissive use the cited definition supports. "Prohibited inference" lists the conclusions the definition does not support. The "review trigger" column states the condition that should route a record to a curator. The table is a policy proposal derived from the cited definitions, not a statement by Mondo.

Relation or annotation	Allowed reuse	Prohibited inference	Review trigger
MONDO:equivalentTo xref (OWL-strength)	Candidate for automatic evidence pooling when the release and artifact are pinned. Mondo defines it as strong equivalence (mondo.readthedocs.io)	That the pairing is clinically interchangeable; that it survives every export (the OBO edition omits the axioms (mondo.monarchinitiative.org))	Source or target term obsoleted, renamed or re-defined in a later release
skos:exactMatch (SSSOM)	Pooling at disease level where a curator-reviewed justification is recorded; the Mondo case study defines it as a source term "conceptually identical to a term in Mondo" (mapping-commons.github.io)	That it equals MONDO:equivalentTo, which the FAQ calls much stronger (mondo.monarchinitiative.org); transitive chaining across three or more vocabularies without checking each hop	Any one-to-many fan-out, or a missing justification or reviewer
skos:closeMatch	Display and search suggestion only	Pooling evidence; the SSSOM guidance advises avoiding the category in practice (mapping-commons.github.io)	Always route to curation before reuse

Relation or annotation	Allowed reuse	Prohibited inference	Review trigger
skos:broadMatch or narrowMatch	Keep as a directional association (for example, roll evidence up to the broader disease)	Traversing in both directions. In SSSOM, broad and narrow are inverse categories (mapping-commons.github.io) and the SKOS inverse relation is defined between the two properties (www.w3.org), but neither is transitive (www.w3.org)	Direction ambiguity: the Mondo case study glosses narrowMatch from the external term's side (mapping-commons.github.io)
MONDO:relatedTo or skos:relatedMatch	Navigation and literature discovery. The editors guide uses relatedTo "when a term is not equivalent but similar" (mondo.readthedocs.io)	Any evidence pooling	None required if never used for aggregation
Bare database xref	Lookup hint	Equivalence of any kind (pmc.ncbi.nlm.nih.gov)	Treat as unmapped until a qualifying relation is found
Exact synonym label	Candidate generation for curators	Identity of two diseases; text matching can mislead (www.medrxiv.org)	Any label-only match, especially for eponymous syndromes
Term replaced by (obsoleted term)	Re-point a stale identifier for review; Mondo adds this annotation holding the replacement CURIE (mondo.readthedocs.io)	Biological equivalence of old and new concept	Every use; replacement scope can widen or narrow
Consider (obsoleted term)	Suggest candidate targets to a human; OBO Format says the replacement needs to be "looked at carefully by a human expert" (owlcollab.github.io)	Automatic substitution	Always

The asymmetry of the middle rows matters most. Exact matches are symmetric, so evidence can flow either way; broad and narrow matches are directional, and rolling evidence up from a narrower disease to a broader one does not authorize pushing broad-disease evidence down. Transitivity belongs to skos:exactMatch alone (www.w3.org), and a chain of exact matches inherits the weakest curation of any hop, so the ledger records every hop with its own version.

Directionality and the SSSOM subject and object roles

SSSOM defines the broadMatch predicate so that "the object is conceptually broader than the subject" (mapping-commons.github.io), and narrowMatch as the inverse, "the object is conceptually narrower than the subject" (mapping-commons.github.io). The Mondo mapping case study glosses its own narrow category from the external term's perspective: "a term in an external source is conceptually narrower to a term in Mondo" (mapping-commons.github.io). The case study's glosses and the SSSOM definitions are written from different vantage points. A pipeline must therefore read the subject and object columns of the pinned mapping file instead of assuming a direction from a prose description.

Disease entities, groupings and phenotypic features

Mondo's FAQ is explicit that a grouping term is not a disease entity: its example is that Parkinson's disease "is NOT a disease entity" in the sense of a single specific condition, and it says current efforts aim to separate disease entities from classes or groupings (mondo.monarchinitiative.org). OMIM phenotypic series are represented as grouping classes (mondo.readthedocs.io). Mondo's exclusion documentation excludes terms that are "not a specific disease" and instead group diseases (mondo.readthedocs.io), and it excludes phenotype terms that represent observable characteristics of an individual (mondo.readthedocs.io). HPO, in contrast, is used to describe human phenotypic abnormalities (obophenotype.github.io), and the Mondo homepage notes that Mondo's development is coordinated with it (mondo.monarchinitiative.org).

For a mapping contract this means a three-way type check before any predicate is examined: is the source concept a disease entity, a group of diseases, or a phenotypic feature? A phenotype term mapped to a disease class, or a group mapped to a single entity, should never be marked as an exact match. The rule should hold even when a source resource conflates categories.

Key Changes: Release Review of September to November 2026

Release-change summary

Table 2 summarizes the published change counts from the GitHub release notes for the three most recent Mondo releases. These counts are the maintainers' headings, reported as published. They count terms in the ontology, not records in any downstream dataset.

Release (published)	New terms	Renamed (excluding obsoleted)	Obsoleted with replacement	Obsoleted without replacement
v2026-09-01 (1 Sep 2026) (api.github.com)	9 (api.github.com)	4 (api.github.com)	2 (api.github.com)	0
v2026-10-06 (6 Oct 2026) (api.github.com)	40 (github.com)	27 (github.com)	13 (github.com)	2 (github.com)
v2026-11-03 (3 Nov 2026) (api.github.com)	39 (github.com)	265 (github.com)	36 (github.com)	0 (github.com)

The same notes also report definition additions and changes and obsoletion candidates. For v2026-10-06 they list 18 text definitions added, 61 changed, 4 new obsoletion candidates and 2 terms removed from candidate status (github.com); for v2026-09-01, 10 added, 19 changed and 6 new candidates (api.github.com); for v2026-11-03, 24 added, 40 changed and 3 new candidates (github.com). The long lists of new and renamed terms are in the tagged notes and the release-diff files, and readers should open those for the full lists.

Table 2 shows a roughly ten-fold jump in renamings between October and November. The 265 renamings in v2026-11-03 include at least one rename that follows an X-linked naming pattern. One listed example renames an X-linked condition to "intellectual developmental disorder, X-linked, syndromic, Cabezas type" (github.com). The release notes do not show whether a rename of this kind changes only the label, so it is not safe to

assume label-only for any row, because the release notes list old and new labels without classifying each rename as label-only or semantic. Reviewing a batch rename therefore means diffing the definition, the parents and the cross-references for each affected identifier, not trusting the label diff.

What the two obsoletion categories mean for review

Mondo never deletes identifiers. Its FAQ says "Mondo terms are never deleted", so every identifier ever created remains in the ontology as an obsolete class (mondo.monarchinitiative.org). Obsoletion is announced in advance: users "will have at least 2 months" to review obsoletion candidates (mondo.monarchinitiative.org). In the editors guide, an obsoleted term receives a "term replaced by" annotation when a replacement exists (mondo.readthedocs.io), a "consider" annotation when no single replacement exists (mondo.readthedocs.io), and owl:deprecated set to true (mondo.readthedocs.io). Mondo adds that a term cannot carry both annotations (mondo.readthedocs.io), and that obsolete terms should never be equivalent to anything (mondo.readthedocs.io).

The two published categories therefore map to two distinct review outcomes. "Obsoleted with replacement" supports a re-point proposal that a curator confirms. "Obsoleted without replacement" means the record cannot be re-pointed automatically; at best a consider pointer suggests where to look. The OBO term-stability principle supplies the underlying rule: a term must always denote the same thing, and when its referents change a new term with a new IRI must be created (obofoundry.org). Mondo's curation guide applies the rule: if the concept changes, the identifier is obsoleted and a new concept is created (mondo.readthedocs.io). A rename that keeps the identifier is, by that policy, a change of label or presentation and not of concept, although the policy depends on curators applying it correctly, which is why a spot check of high-value diseases remains prudent.

Worked release-change cases

Table 3 lists cases taken from the published notes and the Ontology Lookup Service (OLS) records for two obsoleted terms. The "decision" column is this report's proposed routing under the matrix above, not a Mondo ruling.

Case	What the release record shows	Change type	Proposed routing
MONDO:0100351 into **MONDO:0012953** (v2026-10-06)	The obsolete term is replaced by MONDO:0012953 (github.com). In the same release MONDO:0012953 was renamed from a "colorectal cancer, susceptibility to, 10" label to "POLD1-related polyposis and colorectal cancer syndrome" (github.com). OLS shows a merge reason and the replacement pointer (www.ebi.ac.uk)	Merge into a renamed term	Re-point; check that evidence labelled with the old susceptibility wording still fits the surviving concept
MONDO:0012144 (v2026-10-06)	Listed as "obsolete Waardenburg syndrome type 2D" without replacement (github.com). OLS shows no replacement but a consider pointer to MONDO:0019517 (www.ebi.ac.uk)	Obsolete without replacement	Do not re-point; send to curation with the consider candidate
MONDO:0008090 (v2026-10-06)	Renamed from "cyclic hematopoiesis" to "cyclic neutropenia" (github.com)	Rename that changes the named condition	Hold for review; the notes do not say whether the change is semantic

Case	What the release record shows	Change type	Proposed routing
MONDO:0003707 (v2026-10-06)	Renamed from "distal biliary tract carcinoma" to "distal cholangiocarcinoma" (github.com)	Rename	Likely label update; confirm parents and xrefs are unchanged
MONDO:0034021 (v2026-10-06)	Obsolete term replaced by the term labelled "Ehlers-Danlos syndrome, spondylodysplastic type" (github.com)	Obsolete with replacement	Re-point after check that the evidence refers to the same subtype
MONDO:0008002 (v2026-10-06)	Obsolete term replaced by a term labelled "mirror movements 1 and/or agenesis of the corpus callosum" (github.com)	Obsolete with replacement whose label appears broader	Review for granularity; a wider replacement should not silently absorb narrower evidence
MONDO:0001213 and a second obsolete term (v2026-11-03)	"obsolete serous glue ear" is listed with a replacement of otitis media with effusion, and a second obsolete term points to the same target (github.com)	Two-to-one merge	Re-point; count unique diseases before and after
MONDO:0016029 (v2026-11-03)	Obsolete esthesioneuroblastoma is listed with the replacement "olfactory neuroblastoma" (github.com)	Obsolete with replacement	Re-point; probable terminology change, confirm by definition
MONDO:0006008 (v2026-11-03)	Renamed from "vestibular neuronitis" to "vestibular neuritis" (github.com)	Rename	Label update; no re-point needed if identifier and parents are stable

Three lessons follow. A merge can combine an obsoletion and a rename in the same release, so a join keyed on the old label and identifier fails twice. An obsoletion without replacement can still carry a consider pointer, consistent with Mondo's rule that consider is used when no single replacement exists (mondo.readthedocs.io). And two obsolete terms merge into one surviving identifier, so unique diseases fall even though evidence records do not.

Whether a rename is only a label change cannot be read from the notes, which give old and new labels. A defensible check compares definition text, parent classes and cross-references across two pinned releases, using the release-diff tables that accompany each tag. The assets of v2026-10-06 include mondo_release_diff_changed_terms.tsv and a source-versions.tsv file (api.github.com).

Source Vocabularies: OMIM, Orphanet, MeSH and ICD

OMIM and Orphanet

Mondo includes OMIM entries that represent disease concepts, and excludes entries that represent genes or loci (mondo.readthedocs.io). OMIM phenotypic series become grouping classes (mondo.readthedocs.io). A mapping ledger should therefore carry the OMIM entry type, because an OMIM number that denotes a gene

has no Mondo disease target by design. Mondo's source page lists the ontologies used as sources or for cross-references, (mondo.monarchinitiative.org). Confirm the licence conditions of the OMIM service, including reuse and redistribution terms for OMIM-derived data, on the OMIM site before storing any bulk extract.

For Orphanet, an ORPHAcode identifies a clinical entity and is described as unique and time-stable (build.fhir.org). Orphanet assigns each entry a classification level of group of disorders, disorder or subtype of disorder (pmc.ncbi.nlm.nih.gov), which means that the granularity of an ORPHAcode must be read before it is compared with a Mondo class. Orphanet's own mappings to ICD-10 use relation codes that are an instructive model for a policy: E (exact) means the two codes "describe the same disease entity", BTNT means the ORPHAcode "has a broader range of application than the associated ICD-10 code", NTBT means it "has a narrower range than the ICD-10 code used to represent it", and ND is used when the alignment cannot be qualified by any of those relationships (pmc.ncbi.nlm.nih.gov). Licence statements for Orphanet data differ across the pages reviewed. An Orphanet-authored post states that its datasets under CC BY 4.0 allow "integration of our data into third-party resources and tools" (network.febs.org), whereas a Health Level Seven (HL7) terminology record describes extraction as "allowed for any non-commercial use" (build.fhir.org). The statements may refer to different products or dates, and a team should confirm the terms that apply to the file it actually downloads.

MeSH, ICD and the Disease Ontology

MeSH (Medical Subject Headings) is the National Library of Medicine's controlled vocabulary used to index PubMed and other databases (www.nlm.nih.gov). NLM's terms for downloaded data state that no charges, usage fees or royalties are paid to NLM (www.nlm.nih.gov) and require an acknowledgement line (www.nlm.nih.gov). MeSH is a heading vocabulary, so a MeSH heading may denote a disease, a group or a concept used for indexing, and the same type check described earlier applies.

ICD (International Classification of Diseases) is a classification, not a disease ontology. The ICD-11 Reference Guide notes that in a classification there are residual categories for other and miscellaneous conditions (icdcdn.who.int). A residual code such as "other specified" has no disease entity behind it and should map to no Mondo class. The same rule explains why Mondo's exclusion documentation drops terms that merely group diseases (mondo.readthedocs.io). The WHO page states that ICD-11 is distributed under a Creative Commons Attribution-NoDerivs 3.0 IGO licence (www.who.int); confirm the separate ICD-11 terms of use before storing any bulk extract. On coverage, Mondo's ingest report counts exact, broad and narrow mappings together as mapped (monarch-initiative.github.io), so a coverage percentage for ICD-10 in that report is not an exact-match percentage and should be read with that definition in mind. The FAQ lists reasons a source term may be absent from Mondo, including that the term is new and the pipeline has not captured it (mondo.monarchinitiative.org), or that a curation mistake excluded it (mondo.monarchinitiative.org). A missing mapping is therefore a coverage gap and not proof that no corresponding disease exists.

Choosing and Freezing a Release Artifact

Editions and what survives export

The Mondo download page lists the complete OWL edition as the "complete ontology, plus inter-ontology equivalence axioms" (mondo.monarchinitiative.org). It lists the OBO edition as omitting equivalence axioms, with xrefs usable as a proxy (mondo.monarchinitiative.org). The JSON edition is described as equivalent to the OWL edition (mondo.monarchinitiative.org), and a rare-disease subset is available (mondo.monarchinitiative.org). The Mondo homepage describes an edition that uses OWL equivalence axioms directly in the ontology (mondo.monarchinitiative.org). The practical consequence is stark: the same release can yield different logical content depending on the file, so the contract must name the artifact, not only the release.

Mondo mappings are also published in SSSOM format. The Mondo case study confirms that Mondo uses SSSOM to represent mappings (mapping-commons.github.io), and the SSSOM paper reports that Mondo exports SSSOM tables in its release pipeline (pmc.ncbi.nlm.nih.gov). The Mondo FAQ, however, still says "we will soon be releasing these using a new format" (mondo.monarchinitiative.org), wording that appears dated relative to the later sources rather than a direct contradiction. The discrepancy is worth recording in a ledger note and resolving by inspecting the pinned release assets directly.

Pin, hash and verify

The latest-release URL is mutable; at the time of writing it resolved to v2026-11-03, published 3 November 2026 (api.github.com), and the pointer moves with every release. The OBO Foundry versioning principle says the content of official release files must not be changed (obofoundry.org), so a tag-pinned artifact is a stable input. The GitHub API listed v2026-10-06 with mondo.json at 107,871,006 bytes among its assets (api.github.com), and each asset carries a sha256 digest. A frozen manifest should record the tag, the asset name, the byte size, the digest, the retrieval date, and the licence. Mondo's repository carries the Creative Commons Attribution 4.0 International licence text (raw.githubusercontent.com), and the download page states CC BY 4.0 (mondo.monarchinitiative.org). Third-party vocabularies keep their own terms.

Asset counts can differ between a release page summary and the API listing, so the ledger should state which listing it used and why; a digest comparison of the files actually downloaded settles the question better than a count.

Timing also matters. The FAQ states that Mondo is released around the first of each month (mondo.monarchinitiative.org) and appears in OLS within 7 business days (mondo.monarchinitiative.org), and the download page says OLS is uploaded approximately one week after the release (mondo.monarchinitiative.org). The ledger should record which service resolved each identifier.

Implementation Considerations and Process Changes

The mapping-policy ledger

The SSSOM specification makes justification mandatory in spirit: "there can be no mapping without some form of justification" (mapping-commons.github.io), and an implementation must not allow a mapping object that lacks the required slots (mapping-commons.github.io). It records version information for the source (mapping-commons.github.io) and for the mapping set (mapping-commons.github.io), an optional confidence that defaults to 1.0 when unspecified (mapping-commons.github.io), reviewer identity (mapping-commons.github.io), and justification types such as manual curation by domain expertise (mapping-commons.github.io). The SSSOM authors stress that mappings often lack the metadata needed to interpret them (pmc.ncbi.nlm.nih.gov), and that understanding how a mapping was created is crucial to interpreting it (pmc.ncbi.nlm.nih.gov). A project ledger should extend these slots with the fields below.

- **Source concept:** identifier, label, vocabulary, entry type (disease, group, gene, phenotype) and the source release or date.
- **Destination concept:** Mondo identifier, label at the pinned release, and obsolescence status.
- **Relation:** the exact predicate (equivalentTo, exactMatch, broadMatch, narrowMatch, relatedMatch), with direction.
- **Assertion source:** who asserted the mapping (Mondo curators, a vocabulary owner, or the project's own curators) and the justification type.
- **Source and mapping versions:** the source version, the Mondo tag and artifact, and the mapping-set version.
- **Evidence:** the supporting definition, cross-reference set or published rationale, and the confidence value.
- **Review decision:** accept for pooling, keep as directional association, or curate, with reviewer, date and reason.

The Mondo process behind its own mappings supports this level of detail. The 2022 preprint describes a computational strategy that predicts equivalence from labels, synonyms, cross-references, graph structure and priors (www.medrxiv.org), and the 2025 paper says that a proposed mapping from the lexical matcher is reviewed by a curator before being added (escholarship.org). The FAQ states that the semi-automated mapping is always manually evaluated by a curator (mondo.monarchinitiative.org). Equivalence in Mondo is normally one to one, except in "proxy merge" cases in which Mondo decides that two concepts in an external resource mean the same thing (mondo.monarchinitiative.org); a preferred-external annotation then marks the preferred term among the merged ones (mondo.readthedocs.io). Source-to-Mondo cardinality can therefore legitimately be many-to-one, and the ledger must allow it.

Acceptance tests and review routing

A release upgrade should run a fixed set of tests before any downstream join is refreshed.

- **Obsolescence test:** any ledger destination whose identifier is now obsolete is blocked from automatic pooling. If a replacement annotation exists, the record is queued as "re-point proposal"; if only a consider pointer exists, it is queued as "curate".
- **Merge test:** where two source identifiers now resolve to one destination, report unique diseases and evidence records separately and flag the pair for a decision on whether the evidence is additive.
- **Rename test:** for each renamed destination, diff definition, parents and cross-references against the previous pinned tag; escalate any change beyond the label.
- **Type test:** confirm that no source concept typed as a gene, group or phenotype is mapped by an exact relation.
- **Direction test:** reject any broad or narrow mapping that is traversed against its stored direction.
- **Reasoner test:** if exact matches are loaded as OWL equivalences, run a consistency check, since loose matches can create contradictions on merge, as the FAQ warns (mondo.monarchinitiative.org).
- **Notice test:** compare announced obsoletion candidates with the ledger early, within the two-month window (mondo.monarchinitiative.org).

Review routing follows from the outcomes. Records passing all tests proceed. Records with a re-point proposal go to a domain curator who confirms that the replacement matches the evidence's disease granularity. Records without a replacement, direction failures and type failures go to a senior curator or to the ontology's issue tracker. None of these steps replaces clinical judgment, and the output must not be used as diagnostic or eligibility guidance.

Quantifying join effects (Hypothetical Example)

These numbers are hypothetical and illustrate arithmetic only. Suppose a reader's indication dataset has 1,000 evidence records spanning 420 distinct source disease identifiers, and the join to Mondo at the v2026-10-06 pin yields 360 accepted identifiers covering 810 records, 40 ambiguous identifiers covering 120 records, and 20 unmapped identifiers covering 70 records. Accepted identifiers are 85.7 percent of the 420 identifiers (360 of 420) but 81.0 percent of the 1,000 records, because accepted diseases average 2.25 records each. After moving to the next tag, suppose 6 accepted identifiers now point at obsolete destinations and cover 21 records. Holding those for review leaves 354 accepted identifiers (84.3 percent) and 789 accepted records (78.9 percent). If two of the six resolve to the same replacement, unique destination diseases fall by one while the record count is unchanged.

The change in accepted share is a mapping effect, calculated separately from any change in the underlying evidence, and a fall in unique diseases from a merge is not a loss of evidence. Report both denominators.

Data Analysis and Evidence

Ontology scale: why counts disagree

Quoted sizes of Mondo differ because they count different things at different dates. The 2022 preprint reported "over 20,000 disease classes" (www.medrxiv.org). The 2025 Genetics paper reports more than 25,000 disease concepts reconciled from 21 resources (escholarship.org), counted in version v2025-04-01 (escholarship.org), and its table caption states that "Mappings correspond to the number of exact mappings from the Mondo"

disease branch to a resource (escholarship.org). The Mondo homepage, read on 28 November 2026 without a stated release date, shows a total number of diseases of 29,315 (mondo.monarchinitiative.org). The OLS programming interface reported 63,573 terms for version 2026-11-03 (www.ebi.ac.uk), which is evidently a count of all classes rather than diseases. These four numbers are not contradictory, but none is a release-specific disease total for the tag a reader has pinned. A report should therefore count from the pinned artifact using a stated definition (for example, non-obsolete descendants of the disease branch) rather than quote a homepage figure.

Professional reuse evidence

The 2025 paper states that ClinGen is using 1,297 Mondo terms to standardize disease terms (escholarship.org), that Open Targets links its content to 9,815 diseases in Mondo (escholarship.org), and that the National Organization for Rare Disorders (NORD) presents 10,117 rare diseases in Mondo (escholarship.org). It also says ClinGen and MedGen use Mondo as a central authority for defining and curating disease entities (escholarship.org). These are self-reported adoption counts from the Mondo authors at the time of writing. They show reuse, not endorsement of any particular mapping rule.

Independent pages support the reuse picture with qualifications. ClinGen's page says MONDO identifiers are required in its Gene and Variant Curation Interfaces (clinicalgenome.org), but the video was recorded on 10 August 2017 (clinicalgenome.org), so current interface practice should be verified. ClinGen's gene tracker procedure prompts curators to provide a MONDO identifier for the disease entity (gene-tracker.clinicalgenome.org). NCBI's 2020 announcement says ClinVar and the Genetic Testing Registry accept submissions that use MONDO identifiers (ncbiinsights.ncbi.nlm.nih.gov), and the MedGen help page says concepts from OMIM phenotypic series are integrated as part of releases from Mondo (www.ncbi.nlm.nih.gov). Open Targets documents the Experimental Factor Ontology (EFO) as its scaffold for the disease or phenotype entity rather than Mondo itself (platform-docs.opentargets.org), while also stating that the Platform ingests the Monarch merged disease ontology for clinical signs and symptoms (platform-docs.opentargets.org). An Open Targets blog on ClinGen curations reported that, of 493 MONDO terms in the gene validity file, only 94 were already part of EFO (blog.opentargets.org), and an earlier post described EFO working with Mondo on a merged classification (blog.opentargets.org). The Global Alliance for Genomics and Health (GA4GH) Phenopacket documentation lists Mondo among recommended ontologies (phenopacket-schema.readthedocs.io) and recommends OMIM for rare diseases as one option (phenopacket-schema.readthedocs.io). Taken together, the evidence shows that EFO-based and MONDO-based platforms coexist, so crosswalks between them are routine and must follow the same predicate discipline.

Frequently Asked Questions (FAQs)

How do I map disease terms to Mondo? Start from a pinned release and the SSSOM mapping set, type each source concept (disease, group, gene, phenotype), look up an existing curated mapping with a recorded relation, and send everything else to curation. Use label matching only to propose candidates, because text matching can be misleading (www.medrxiv.org).

What is the difference between MONDO:equivalentTo and skos:exactMatch? The Mondo FAQ says they conceptually overlap but have entirely different semantics, with equivalentTo much stronger (mondo.monarchinitiative.org).

How do Mondo, OMIM and Orphanet mappings relate? Mondo includes OMIM disease entries and excludes gene or locus entries (mondo.readthedocs.io), and Orphanet classifies entries at group, disorder and subtype levels (pmc.ncbi.nlm.nih.gov). Granularity must be compared before any evidence is pooled.

How does Mondo compare with ICD-10 mapping? ICD is a classification with residual categories (icdcdn.who.int), whereas Mondo targets specific disease entities and excludes grouping terms (mondo.readthedocs.io). Exact matches to ICD-10 should be expected to be incomplete, and broad or narrow relations should be kept directional.

What changed in the latest Mondo release? v2026-11-03 lists 39 new terms, 265 renamings and 36 obsoletions with replacement (github.com). See Table 2.

Do obsolete Mondo identifiers disappear? No. Identifiers are never deleted and remain as obsolete terms (mondo.monarchinitiative.org).

Is an obsolete term's replacement biologically equivalent? Not automatically. Mondo reserves "consider" for cases without one replacement (mondo.readthedocs.io), and OBO Format says a consider candidate must be inspected by a human expert (owlcollab.github.io). A replacement annotation is also a curator's pointer and needs a granularity check.

Implications and Future Directions

A disease-mapping table is a versioned dataset with its own change log. A release review should be a scheduled job, tied to the roughly monthly cadence (mondo.monarchinitiative.org), that compares the ledger against the new tag, produces Table 2-style counts for the reader's own records, and routes exceptions. Teams integrating EFO, ClinVar or Open Targets content should expect identifiers from several resources to converge on one Mondo class, increasing many-to-one merges.

The ledger approach also supports a policy on direction. Because skos broad and narrow relations are inverse but not transitive (www.w3.org), roll-ups should be implemented as explicit one-way queries, with no automatic downward propagation of evidence. A policy should further state a cutoff for confidence values, remembering that SSSOM defaults an unspecified confidence to 1.0 (mapping-commons.github.io); a ledger that treats a default as a measured score would overstate certainty.

Several uncertainties remain. The release notes do not classify renames as label-only or semantic. The FAQ wording on SSSOM appears dated relative to the later papers, so check the release assets directly, and confirm OMIM licence terms before bulk storage. Mondo's work to separate disease entities from groupings is ongoing (mondo.monarchinitiative.org), so expect further structural change. This report measures no dataset's real join rates, and the hypothetical example is illustrative until computed on the reader's own data. The practical framing is to treat the work as a documentation and governance exercise: the value lies in the written, reviewable decision, not in a single crosswalk file.

Conclusion

Mondo disease mappings are useful when the relation behind each match is explicit. A cross-reference is a lookup hint, an exact synonym is a label, skos:exactMatch is a curated claim of conceptual identity, and MONDO:equivalentTo is a strong logical axiom that Mondo says is much stronger than exactMatch. Evidence pooling should rest on the stronger relations, with broad and narrow matches kept as directional associations and everything else sent to curation.

The October and November 2026 release records show why a release-specific review matters. Between v2026-10-06 and v2026-11-03 the published count of renamings rose from 27 to 265 and obsoletions with replacement rose from 13 to 36, while v2026-11-03 listed none without replacement. None of those counts says how many records in a given dataset change. Worked cases show merges into renamed terms, an obsoletion with only a consider pointer, and at least one rename that changes the named condition. Each requires a recorded decision.

The practical contract is short: pin the tag and artifact, keep the ledger fields for source, destination, relation, assertion source, versions, evidence and review decision, report unique diseases and evidence records separately, and treat replacements as proposals. The result supports research decisions about indications. It is not clinical coding, diagnosis or trial-eligibility advice.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://mondo.monarchinitiative.org
2. https://escholarship.org/content/qt7d51t1mb/qt7d51t1mb.pdf
3. https://mondo.monarchinitiative.org/pages/faq
4. https://api.github.com/repos/monarch-initiative/mondo/releases/latest
5. https://github.com/monarch-initiative/mondo/releases/tag/v2026-11-03
6. https://github.com/monarch-initiative/mondo/releases/tag/v2026-10-06
7. https://obofoundry.org/principles/fp-004-versioning.html
8. https://mondo.readthedocs.io/en/latest/editors-guide/merging-and-obsoleting
9. https://obophenotype.github.io/human-phenotype-ontology
10. https://mondo.readthedocs.io/en/latest/editors-guide/exclusion-reasons
11. https://clinicalgenome.org/tools/educational-resources/materials/monarch-disease-ontology-mondo
12. https://ncbiinsights.ncbi.nlm.nih.gov/2020/05/12/mondo-ids-clinvar-gtr
13. https://www.medrxiv.org/content/medrxiv/early/2022/05/03/2022.04.13.22273750.full.pdf
14. https://mondo.readthedocs.io/en/latest
15. https://pmc.ncbi.nlm.nih.gov/articles/PMC9216545
16. https://www.w3.org/TR/skos-reference
17. https://mapping-commons.github.io/sssom/dev/mapping-predicates
18. https://mondo.readthedocs.io/en/latest/editors-guide/f-entities

19. https://www.w3.org/TR/owl2-primer
20. https://mondo.monarchinitiative.org/pages/download
21. https://mapping-commons.github.io/rda-fair-mappings/case-study/case-study-mondo
22. https://owlcollab.github.io/oboformat/doc/GO.format.obo-1_4.html
23. https://mondo.readthedocs.io/en/latest/editors-guide/patterns/OMIM_phenotypic_series
24. https://api.github.com/repos/monarch-initiative/mondo/releases/tags/v2026-09-01
25. https://api.github.com/repos/monarch-initiative/mondo/releases/tags/v2026-10-06
26. https://obofoundry.org/principles/fp-019-term-stability.html
27. https://mondo.readthedocs.io/en/latest/editors-guide/using-mondo-for-curation
28. https://www.ebi.ac.uk/ols4/api/ontologies/mondo/terms/http%253A%252F%252Fpurl.obolibrary.org%252Fobo%252FMONDO_0100351
29. https://www.ebi.ac.uk/ols4/api/ontologies/mondo/terms/http%253A%252F%252Fpurl.obolibrary.org%252Fobo%252FMONDO_0012144
30. https://mondo.readthedocs.io/en/latest/editors-guide/OMIM-representation
31. https://mondo.monarchinitiative.org/pages/sources
32. https://build.fhir.org/ig/HL7/UTG/en/ECS_MR-orphanet.html
33. https://pmc.ncbi.nlm.nih.gov/articles/PMC10476382
34. https://network.febs.org/posts/orphadata-data-and-services-to-harness-the-power-of-the-orphanet-knowledge-base
35. https://www.nlm.nih.gov/databases/download/mesh.html
36. https://www.nlm.nih.gov/databases/download/terms_and_conditions.html
37. https://icdcdn.who.int/icd11referenceguide/en/refguide.pdf
38. https://www.who.int/standards/classifications/classification-of-diseases
39. https://monarch-initiative.github.io/mondo-ingest/reports/unmapped
40. https://raw.githubusercontent.com/monarch-initiative/mondo/master/LICENSE
41. https://mapping-commons.github.io/sssom/dev/spec-model
42. https://mapping-commons.github.io/sssom/dev/subject_source_version
43. https://mapping-commons.github.io/sssom/dev/mapping_set_version
44. https://mapping-commons.github.io/sssom/dev/confidence
45. https://mapping-commons.github.io/sssom/dev/reviewer_id
46. https://mapping-commons.github.io/sssom/dev/mapping-justifications
47. https://www.ebi.ac.uk/ols4/api/ontologies/mondo
48. https://gene-tracker.clinicalgenome.org/files/SOP_V1.pdf
49. https://www.ncbi.nlm.nih.gov/medgen/docs/help
50. https://platform-docs.opentargets.org/disease-or-phenotype
51. https://platform-docs.opentargets.org/disease-or-phenotype/clinical-signs-and-symptoms
52. https://blog.opentargets.org/introducing-clingen-gene-validity-curations
53. https://blog.opentargets.org/efo3-a-community-driven-ontology-to-advance-clinical-discoveries

54. https://phenopacket-schema.readthedocs.io/en/latest/recommended-ontologies.html

55. https://phenopacket-schema.readthedocs.io/en/latest/disease.html

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.