Test and validation plan with release 1.0 results

Acceptance criteria, synthetic test cases and what has and has not been validated.

1. Acceptance criteria

Criterion	Result	Evidence
All specified patient forms are functional	Pass	All 14 record tabs open, save and reload (automated browser test).
All major organ systems represented	Pass	22 systems, 232 template findings, plus free search under each system.
HPO IDs and labels validated against a documented release	Pass	All template and plain-language terms resolved against release 2025-01-16 at build time.
Negative findings separate from missing, unknown, unassessed	Pass	Six statuses plus “not recorded”; exported as excluded in Phenopackets.
Age of onset and provenance retained	Pass	Per observation and per annotation.
AI-extracted findings require human review	Pass	Suggestions stay Pending; only Accept creates an observation (tested).
Records export in documented formats	Pass	Seven patient exports and four cohort exports produced in the test run.
Data integrity across modules	Pass (browser edition)	Single encrypted document per patient; reload test passed.
Permissions and audit logging	Partly	Role modes and audit trail work; individual accounts need the server edition.
Synthetic end-to-end cases pass	Pass	Six cases: paediatric, adult, family, incomplete, longitudinal, multi-phenotype.
Security controls and backup / restore tested	Partly	Encryption, lockout and backup export tested; no penetration test.
Ontology updates do not silently alter history	Pass	Superseded identifier detected; replacement applied only on request; both kept (tested).
Extensible without major change	Pass	Forms are schema-driven; new fields are one line each.

2. Synthetic test cases (built into the tool)

Case	Content	What it exercises

A — paediatric	Seizures, developmental delay; referral letter; exome result	Extraction with negation, uncertainty and family context; accept / reject; ranking flags a non-specific phenotype
B — adult	Marfan-like features; affected father	Ranking returns FBN1 / TGFBR genes; panels TBO-GP-010 and TBO-GP-011
C — family-based	Hearing loss; consanguineous parents; ten relatives	Three-generation pedigree, double line, uncertain relative not drawn, PED export
D — incomplete	Neonate, one unmapped observation, “not assessed” entry, old HPO identifier	Unmapped observation kept; ontology check and reviewed migration
E — longitudinal	Muscle weakness; three CK results; four follow-ups	Automatic high flags; follow-up due count
F — multi-phenotype	Breast cancer recorded twice; suspected finding; VUS with conflicting evidence; research consent withdrawn	Duplicate flag; suspected kept apart; research export blocked

3. Text-extraction check

Built-in reference set: 15 synthetic sentences, 23 findings. Result for extractor 1.0: precision 96 %, recall 96 %, F1 96 %, context (present / absent / suspected / family) 100 % of matched findings. One finding missed, one false positive. Run it again from “Extraction check” after any change.

Limits: the sentences are short and clean. Before clinical use, build a reference set of at least 200 real, manually reviewed notes per language and specialty and report precision, recall, F1 and mapping accuracy per organ system, with an error analysis.

4. Still to test

- Usability testing with clinicians on real workflows.
- FHIR bundle against a validator and chosen profiles.
- Phenopacket import into Exomiser and LIRICAL with real variant files.
- Ranking benchmark on solved cases (rank of the causal gene).
- Backup restore drill on a second device; accessibility audit (keyboard and screen reader).
- Server edition: unit, integration and load tests; penetration test.

© Tribro-Omics · Test and validation plan with release 1.0 results · Version 1.0 · 11 November 2026