# The Nullary Analytics Suite: Coverage, Tractability, and Cross-Modality Analysis over Measured Negative Results

Methods & validation

Nullary Team

*nullary.ai · Technical Report · 25 May 2026*

**Abstract**

The premium Analytics suite turns Nullary's measured-negative-results layer (122M findings) into four analytics: **coverage maps**, **failure timelines**, a **tractability score**, and **cross-modality analysis**. This report documents their methods and validation, and is deliberately honest about where each is strong or modest. Coverage and timelines are descriptive views over the layer and are exact as of ingestion. The tractability score (built from how heavily a target has been pursued, the best activity achieved against it, and its failure rate) predicts clinical progression at ROC-AUC 0.79. We check it against Open Targets two ways, which are *not* the same claim. It *agrees* with Open Targets clinical tractability (AUC 0.81, 86% agreement) — a consistency check, since that bucket is itself ChEMBL-derived. The genuine external test is that it predicts Open Targets' *structure-derived* druggability (independent of our bioactivity data) at AUC 0.78. Its lift over a trivial popularity baseline is small but statistically significant (full-model bootstrap CIs exclude zero), and CRISPR essentiality is near chance (0.55) for tractability. Cross-modality analysis finds CRISPR essentiality is **largely orthogonal** to chemical druggability (Spearman -0.08); the strong relationships sit within the chemistry/clinical axis (best activity vs clinical phase +0.41). Underpinning all of it, the negative signal reproduces across independent data sources (CRISPR cross-source concordance 98%).

## 1. The data layer

The suite is computed over Nullary's normalized layer of measured negative results: 122,276,636 findings over 1,621,294 compounds, keyed on target and modality with full provenance (Table 1). Unlike positive-only resources, it records what was *tried and failed*—the substrate for coverage, tractability and exhaustion analytics.

| Modality | Findings | Principal sources |
|---|---|---|
| small molecule | 84.5M | PubChem, ChEMBL |
| CRISPR / functional genomics | 37.6M | BioGRID-ORCS, DepMap |
| clinical trial | 104k | AACT, EU-CTR, drugs@FDA |
| peptide / antibody / other | ~51k | THPdb, TheraSAbDab, PROTAC-DB, … |

**Table 1.** Coverage of the negative-results layer by modality (snapshot).

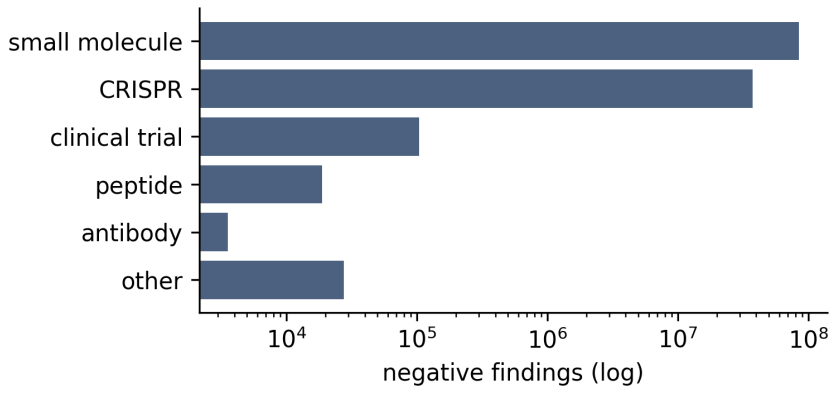


**Figure 1.** Negative findings per modality (log scale).

## 2. Coverage maps and failure timelines

**Coverage maps** answer ‘what has been tried against this target, and how did it fail?’ For any gene or UniProt accession the suite returns the distinct compounds/agents tried-and-failed, broken down by modality and outcome, with a curated-source count (e.g. EGFR: 5,953 distinct compounds across 12,585 negative findings spanning small molecules, PROTACs and peptides). These are exact aggregations over the layer—descriptive, not modelled.

**Failure timelines** place those failures in time, using the dated provenance every finding carries (publication year for bioactivity, trial start/termination dates for clinical records). They show when a target or chemotype was most heavily—and unsuccessfully—pursued. Like coverage maps, they are descriptive surfaces over recorded data and carry no predictive claim.

## 3. Tractability scoring

A tractability score should estimate how amenable a target is to a successful program. We derive candidate features per target from the layer: the breadth of compounds tried (a popularity proxy), the *best activity ever achieved* against it, its small-molecule failure rate, and its CRISPR essentiality. The critical question (and a hazard the companion report flags) is whether these add signal *beyond* the trivial fact that popular targets accrue more drugs.

### 3.1 Calibration against clinical progression

Ground truth: a target’s maximum clinical phase among mechanism-of-action drugs (ChEMBL). Across 2,446 targets, predicting ‘reached clinic’ gives ROC-AUC 0.785 (full feature set). Best single feature is the best activity achieved (0.774). Crucially, the negatives/quality features *without* the popularity proxy reach 0.767—essentially matching popularity-only (0.750)—so the score is not merely counting compounds. CRISPR essentiality alone is near chance (0.55).

### 3.2 Independent validation against Open Targets

ChEMBL phase is partly circular (our features and the label share a source). Open Targets separates clinical tractability (also ChEMBL-derived) from *structure-derived* druggability (PDB pockets/ligands), which is independent of our bioactivity data. Two results (2,436 targets), reported as two distinct claims: **(i) a consistency check** — our label agrees with Open Targets clinical tractability (AUC 0.81, 86% agreement); but that bucket is itself ChEMBL-derived, so this shows reproducibility against a shared-source label, not independence. **(ii) the genuine external validation** — on the *independent* structural-druggability label our features reach AUC 0.778 (quality-only 0.764), so the negatives-based signal tracks an external druggability measure it never saw (Table 2, Fig. 2).

Are the small lifts real? Bootstrapping the AUC differences (5,000 resamples of targets) says yes for the full model: adding the negatives to the popularity proxy beats popularity-only by 0.036 [95% CI 0.023, 0.048] on the ChEMBL label and 0.023 [0.005, 0.042] on the independent structural label — both intervals exclude zero. The negatives *without* the popularity proxy do *not* reliably beat it (CIs straddle zero). So the honest statement is precise: the negatives add a small but statistically reliable increment *on top of* popularity; they augment a popularity score rather than replace it.

| Predicting | popularity-only | quality-only | full |
|---|---|---|---|
| Reached clinic (ChEMBL) | 0.750 | 0.767 | 0.785 |
| Structural druggability (Open Targets, independent) | 0.754 | 0.764 | 0.778 |

**Table 2.** Tractability ROC-AUC by feature set. ‘quality-only’ excludes the popularity proxy. The modest gap (quality vs popularity) is the honest measure of what the negatives add.

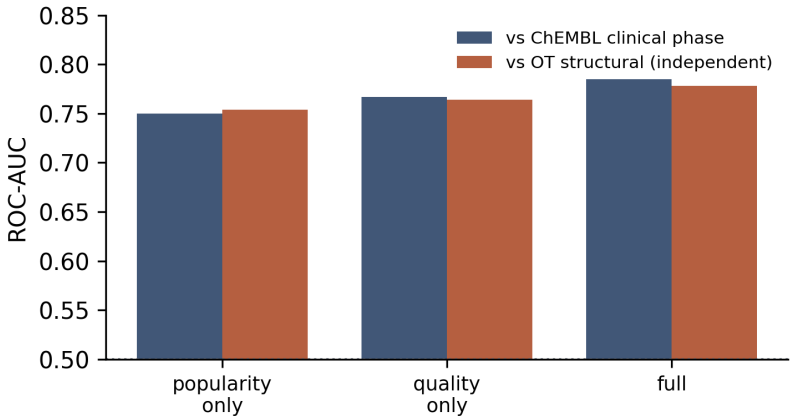


**Figure 2.** Tractability ROC-AUC against ChEMBL clinical phase and against Open Targets' independent structural druggability. The two agree closely; the lift over the popularity baseline is real but small.

**Honest summary.** The tractability score is externally validated and genuinely negatives-grounded—but it is a *modest* discriminator (small lift over popularity, no CRISPR contribution). We position it as a calibrated competitive-landscape / coverage score, not a high-accuracy predictor of program success.

## 4. Cross-modality analysis

Does failure in one modality predict failure in another? Across the same target cohort we correlate CRISPR essentiality, small-molecule druggability and clinical progression (Fig. 3, Table 3). The headline is a *negative* result worth stating plainly: CRISPR essentiality is **largely orthogonal** to chemical and clinical tractability (Spearman -0.08 vs clinical phase, -0.07 vs best activity)—a gene being a dependency does not make it chemically druggable, and vice versa. The strong, actionable relationships live within the chemistry/clinical axis: best activity achieved vs clinical phase +0.41, and failure rate vs clinical phase -0.21.

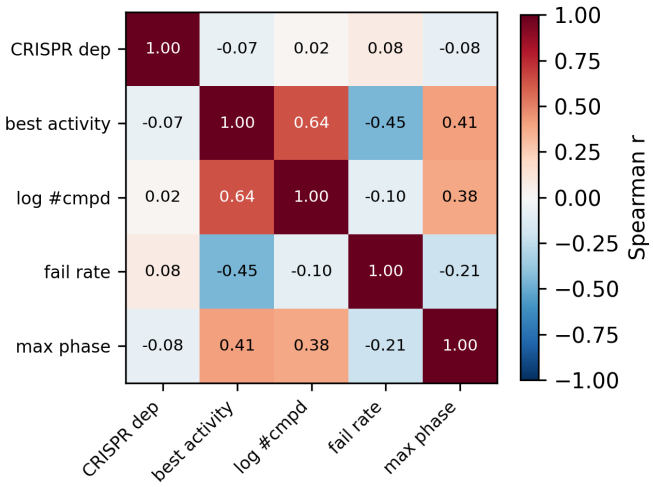


**Figure 3.** Cross-modality Spearman correlations. CRISPR essentiality (top/left) is near-zero with the chemical and clinical axes.

| Relationship | Spearman r |
|---|---|
| best activity ↔ clinical phase | +0.41 |
| failure rate ↔ clinical phase | -0.21 |
| CRISPR essentiality ↔ clinical phase | -0.08 |
| CRISPR essentiality ↔ best activity | -0.07 |

**Table 3.** Key cross-modality correlations. The CRISPR–chemistry pairs are near zero: the modalities carry complementary, not redundant, information.

Two caveats temper the reading: the near-zero CRISPR–chemistry correlation likely reflects *both* genuine biological orthogonality *and* join-level noise from gene-to-target mapping (Limitation 3, below); and it is consistent with Section 3.1, where CRISPR essentiality alone failed (AUC 0.55) to predict clinical progression. For the product this is a feature, framed correctly: cross-modality views let a scientist see that (say) a target is a strong CRISPR dependency yet chemically intractable—*because* the axes are independent. We do not market strong cross-modality predictors, because the data say they largely do not exist.

We tested that claim directly rather than inferring it from the correlation: does *combining* CRISPR essentiality with the chemistry features (via CRISPR×chemistry interaction terms) predict clinical progression better than the additive model? It does not — a likelihood-ratio test on the interactions is non-significant ($p = 0.20$), neither interaction coefficient is significant, and cross-validated AUC is unchanged (+0.0003 [95% CI -0.0014, 0.0019]). The cross-modality *combination* carries no predictive value here over single-modality features — so we offer the cross-modality *view*, not a cross-modality *predictor*.

## 5. Data-quality credential: the negatives are real

Every analytic above rests on the negatives being trustworthy. The genetics and chemistry sides have *separate* credentials. **CRISPR (genetics) — cross-source reproducibility.** For 17,608 genes measured in both DepMap and the independent BioGRID-ORCS collection, essentiality calls (dependency probability ≥ 0.5) agree 98% of the time (cross-prediction AUC 0.99). That overall figure is dominated by the many easy-to-agree non-essential genes, so the telling number is the *essential class* itself: precision 97%, recall 82% (F1 0.89) — agreement is not just an artifact of easy negatives. **Small molecule (chemistry) — decoy-bias awareness.** The companion report shows models trained on these measured inactives beat decoy-trained models on realistic tests, and that decoy validation is systematically over-optimistic — evidence we understand and avoid positive-only benchmark biases. The chemistry side does *not* yet have an analogous cross-source (e.g. ChEMBL–PubChem) concordance check; that is future work.

## 6. Limitations

**(1)** The tractability score's lift over a popularity baseline is small — though statistically significant (full-model bootstrap CI excludes zero, Section 3.2) — so it is a calibrated ranking aid, not a high-accuracy predictor. **(2)** Tractability ground truth (clinical phase; structural druggability) is itself imperfect and time-dependent; and 'reached clinic' is recorded only for pursued programs over a cohort already filtered to targets with ChEMBL activity, so the score is validated on the narrower question 'of popular-enough targets, which reached clinic' rather than tractability at large (a survivorship/popularity bias). **(3)** Cross-modality coverage is uneven—CRISPR is gene-level, chemistry compound-level—and the join is by gene symbol, which dilutes the cross-modality correlation (Section 4). **(4)** Coverage and timelines are descriptive: they report what is recorded, and absence of a failure is not evidence of tractability. **(5)** Findings are auto-extracted; the chemistry side awaits a manual-curation benchmark, while the CRISPR side already has the cross-source check of Section 5. The predictive validation of the underlying activity models, with its controls, is in the companion report [5].

## 7. Availability

Coverage maps and the target-exhaustion index are live via the MCP tool *get_target_landscape* and the REST API. Tractability and cross-modality analytics are computed by the pipelines described here (deterministic, reproducible from the cited public releases: ChEMBL 35, DepMap 26Q1, BioGRID-ORCS, Open Targets 26.03). Contact: nullary.ai.

## References


[1] Mendez D, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* 2019;47(D1):D930–D940.

[2] Tsherniak A, et al. Defining a cancer dependency map. *Cell* 2017;170(3):564–576.

[3] Oughtred R, et al. The BioGRID database and ORCS CRISPR screen repository. *Protein Sci.* 2021;30(1):187–200.

[4] Ochoa D, et al. The next-generation Open Targets Platform. *Nucleic Acids Res.* 2023;51(D1):D1353–D1359.

[5] Nullary Team. Real measured negatives as a substrate for calibrated, target-specific bioactivity prediction. Technical report, 2026.