Mechanism-Based Trial Matching Reveals a 54% Target Alignment Gap in Ovarian Cancer: Quantifying the Precision-Ineligible Majority

Fahad Kiani

April 18, 2026

Abstract

Background: Phase 2 oncology trials fail often. Heterogeneity inside protocol-eligible cohorts dilutes signal when inclusion passes without assurance that tumor biology engages the drug's mechanism of action (MoA). Eligibility tooling rarely reports **mechanism alignment** between tumor profile and trial MoA.

Methods: Patients and trials were represented as **seven-dimensional mechanism vectors** (DDR, MAPK, PI3K, VEGF, HER2, IO, efflux). Mechanism fit was defined as the **magnitude-weighted** score (p * t) / ||t||_2, clipped to [0,1], rather than cosine similarity. Trial-menu discrimination was evaluated on a curated 59-row ovarian cancer menu using a DDR-high reference profile. Matchability prevalence was then estimated on a frozen TCGA-OV cohort by scoring each patient against the prespecified trial menu and classifying strong alignment as `best_fit > 0.5`. Overall survival was analyzed only as a negative control because treatment assignment in TCGA was not mechanism-directed.

Results: On the curated 59-row menu, DDR-directed trials yielded **n = 31** and mean mechanism fit **0.874**, versus **n = 17** and mean fit **0.038** for non-DDR trials (23-fold separation; Delta = 0.836). Eleven additional curated entries collapsed to zero 7D mechanism vectors after sanitization and were retained in the manifest, but not in the binary comparison, because the magnitude-weighted score is undefined for a zero-norm trial vector. In the frozen TCGA-OV branch, **271 of 585 patients (46.3%)** exceeded `best_fit > 0.5`, whereas **314 of 585 (53.7%)** did not. Among the **571** patients with observed overall survival, Cox analysis yielded **HR = 1.122** with **p = 0.288**, consistent with a negative control rather than treatment benefit.

Conclusions: Eligibility-first workflows miss mechanism mismatch unless alignment is measured directly. In this ovarian cancer analysis, explicit mechanism scoring exposed both sharp trial-level discrimination and a majority of patients who remained below a strong-alignment threshold. The near-null TCGA sur-

vival contrast reinforces that retrospective outcome mining is not a substitute for mechanism-enriched enrollment.

---

1. Introduction

1.1 Problem: diluted signal in heterogeneous Phase 2 cohorts

Many Phase 2 programs stop for lack of efficacy. One driver is **heterogeneity among protocol-eligible patients**: criteria satisfied without assurance that tumor biology engages the investigational MoA. That mix widens confidence intervals and can mask an effect visible under mechanism enrichment.

1.2 Eligibility is not mechanism

Eligibility encodes organ function, histology, treatment line, contraindications, and sometimes single-biomarker gates. It does not, by default, yield a scalar **pathway overlap between tumor burden and trial MoA**. Trial lists can be **eligible** and **mechanistically cold** for large fractions; under an explicit alignment rule that defines a **precision-ineligible majority**.

1.3 Contribution: measurement and reporting rules

The missing piece for reporting is **reproducible measurement**: fixed axes, fixed normalization, and explicit non-goals. In particular, TCGA survival is treated here as a negative control rather than as validation of matching efficacy.

1.4 Scope (three items)

1. **Contract:** 7D patient and trial vectors, axis order `PATHWAYS_7D` in `mechanism_fit_ranker.py`.

2. **Fit:** magnitude-weighted (p * t) / ||t||_2 (clipped) instead of cosine for the low-burden false-positive case.

3. **Reporting:** (a) curated-menu discrimination, (b) full-registry inventory diagnostics, and (c) TCGA matchability plus a survival negative control.

---

2. Methods

2.1 Axes and order

Seven axes, fixed order: **DDR, MAPK, PI3K, VEGF, HER2, IO, efflux**. Order is binding for serialization and dot products (`PATHWAYS_7D`). Some trials include an eighth axis, **`rss`**; for 7D patients the ranker projects 8D trial vectors to 7D by dropping `rss`.

2.2 Trial MoA registry

Trial mechanism vectors were stored in a frozen NCT-keyed registry assembled across multiple curation waves. Table 1 summarizes the curated publication menu and the zero-vector subset that required separate accounting.

Hygiene: Non-numeric placeholders (`"No"`, `"N/A"`, etc.) were coerced to **0.0** before vector assembly so that registry-wide scoring used a deterministic numerical contract.

2.3 Patient mechanism vector

Patient vectors used the same seven axes. In the frozen TCGA-OV branch, vectors were constructed from `hrd_proxy`, `brca_somatic`, `tmb`, and `msi_status`. Because the validated cohort file did not contain matched measurements for MAPK, PI3K, VEGF, or HER2 activity, those axes remained 0.0 in the prevalence analysis. The resulting prevalence estimate should therefore be read as conservative with respect to unmeasured non-DDR biology.

2.4 Mechanism fit

Equation: `mechanism_fit(p, t) = clip_[0,1]((p * t) / ||t||_2)`

This is **not** cosine similarity: ||p||_2 does not appear in the denominator. Trials with mass on few MoA-heavy axes should not read as "high fit" when the patient carries only trace mass on those axes.

With eligibility e in [0,1]:

Equation: `combined = 0.7e + 0.3 * mechanism_fit`

Production ranking uses eligibility $>= 0.60$ and mechanism fit $>= 0.30$. The separate mechanism-fit gate means a mechanism-null trial cannot enter the ranked set on eligibility alone. Once both gates are passed, however, the 70/30 blend still favors eligibility: a **0.10** increase in eligibility contributes the same composite score as a **0.233** increase in mechanism fit. The analyses below therefore emphasize mechanism-fit discrimination and TCGA prevalence directly, rather than treating the blended score as an independent biological result.

2.5 Discrimination (DDR reference profile)

Reference 7D: DDR 0.88; MAPK 0.12; PI3K 0.05; VEGF 0.02; HER2, IO, efflux 0. **Labels: DDR-tagged** if `moa_vector["ddr"] > 0.5` after string hygiene; otherwise **non-DDR**.

Curated menu versus binary split: The curated publication menu contains **59** trial-drug annotations. The quoted DDR versus non-DDR comparison uses **48** rows (**31 + 17**). The remaining **11** rows collapse to a **zero 7D mechanism vector** after sanitization and are therefore reported separately rather

than forced into a binary fit contrast with an undefined denominator. For context, the same DDR-high reference profile can also be applied to the full 609-row registry as an inventory-level diagnostic.

2.6 TCGA-OV matchability

Per patient, a 7D vector was scored against the **prespecified trial menu**, and the patient was classified as strongly aligned when **max fit** exceeded `best_fit > 0.5`. In the validated frozen branch, the prevalence cohort contained **n = 585** patients and the survival subset with observed OS contained **n = 571**.

2.7 Survival negative control

High- and low-fit groups were contrasted under usual care. Under non-mechanism-directed treatment assignment, **no material survival difference is expected**. A near-null result therefore functions as a check against retrospective outcome mining, not as evidence that mechanism alignment lacks biological meaning.

---

3. Results

3.1 Registry inventory

After coercion of non-numeric MoA entries, the frozen registry contained **609** NCT-keyed records. This number describes the full mechanism registry and is distinct from the 59-row curated publication menu.

3.2 Discrimination

On the curated 59-row publication menu, DDR-directed trials showed mean mechanism fit **0.874** across **31** rows, versus **0.038** across **17** non-DDR rows, for a 23-fold separation. Eleven additional curated entries collapsed to zero 7D mechanism vectors after sanitization; these rows are retained in the manifest and Table 1, but not assigned to the binary fit contrast because the metric is undefined for a zero-norm trial vector. When the same DDR-high reference profile was applied to the full 609-row registry, the direction of effect was preserved: DDR-tagged trials (**n = 75**) had mean fit **0.821**, whereas non-DDR trials (**n = 534**) had mean fit **0.035**.

3.3 Matchability prevalence

In the frozen TCGA-OV branch, **271 of 585 patients (46.3%)** exceeded `best_fit > 0.5`, whereas **314 of 585 (53.7%)** did not. Under the present framing, that below-threshold majority is the **precision-ineligible majority**: patients who may remain broadly trial-eligible yet do not achieve strong mechanism alignment to the curated menu under the stated rule. Because several

non-DDR axes were unmeasured in the frozen cohort file and therefore fixed to 0.0, this prevalence estimate should be interpreted as conservative.

3.4 Survival negative control

On the **571** TCGA-OV cases with observed OS, the frozen branch yielded **269** matchable versus **302** non-matchable patients, with median OS **1330** versus **1471** days, log-rank $\mathbf{p = 0.2878}$, and Cox $\mathbf{HR = 1.1221}$ ($\mathbf{p = 0.2883}$). The near-null result is the expected outcome in a non-mechanism-enrolled cohort and argues against retrospective survival mining as a substitute for mechanism-guided enrollment.

3.5 Sensitivity analyses

Threshold sweeps are summarized in **Figure S1**, showing the prevalence split across strong-alignment thresholds around the locked `best_fit > 0.5` rule. Additional planned sensitivity analyses include trial-menu sensitivity and pathway ablations (IO and/or efflux dropped).

---

4. Discussion

4.1 Interpretation within scope

Measurement-first: eligibility-first navigation can coexist with a large **mechanism-mismatch** population once fit is measured and reported. Magnitude-weighted fit is a **screening geometry** choice with a documented cosine failure mode on low burden; it is **not** offered here as a validated predictor of response. The central claim is therefore about hidden alignment structure, not about retrospective proof of treatment benefit.

4.2 Survival

The null is informative. Under unenriched treatment assignment, survival separation should be weak or absent even if mechanism mismatch is real. The observed HR near 1 therefore supports a narrower and more useful conclusion: retrospective outcomes alone are a poor benchmark for trial-matching quality when patients were never enrolled by mechanism in the first place.

4.3 Limitations

- **Zero-vector trial entries:** Eleven of the 59 curated rows collapsed to zero 7D mechanism vectors after sanitization. Those trials were not discarded; they were carried forward as an explicit accounting category because the magnitude-weighted score is undefined when the trial vector has zero norm.

- **Conservative TCGA prevalence branch:** The frozen TCGA-OV cohort supported DDR/BRCA/TMB/MSI-derived features, but not matched measurements for several non-DDR axes. Prevalence estimates therefore understate alignment opportunities that would only be visible with broader pathway measurements.

- **Frozen trial-annotation contract:** The trial MoA registry was assembled from mixed automated and manual curation. This study evaluates alignment under that frozen registry; it does not prove that all ovarian cancer trials have complete or final mechanism annotation.

- **Retrospective cohort design:** TCGA was used to estimate matchability prevalence and to supply a survival negative control. It cannot establish the clinical effect of enrolling patients prospectively by mechanism score.

---

5. Conclusions

A **precision-ineligible majority** becomes measurable once mechanism alignment is scored explicitly rather than inferred from eligibility alone. In this ovarian cancer analysis, a DDR-high reference profile sharply separated DDR-directed from non-DDR trials, while the frozen TCGA-OV branch showed that most patients still fell below a strong-alignment threshold. The near-null survival control underscores the practical implication: if enrollment is not mechanism-guided, retrospective outcomes will not reveal the alignment gap that the score exposes.

Data and code availability

Frozen analysis artifacts supporting this manuscript are archived in the `crispro-ai/Publications 02-trial-matching` package together with the locked subset and TCGA freeze artifacts used for the present build. Core scoring code and the frozen trial mechanism registry are stored in the associated CrisPRO application repository.

Figures

Tables

Table 1. Curated publication menu provenance

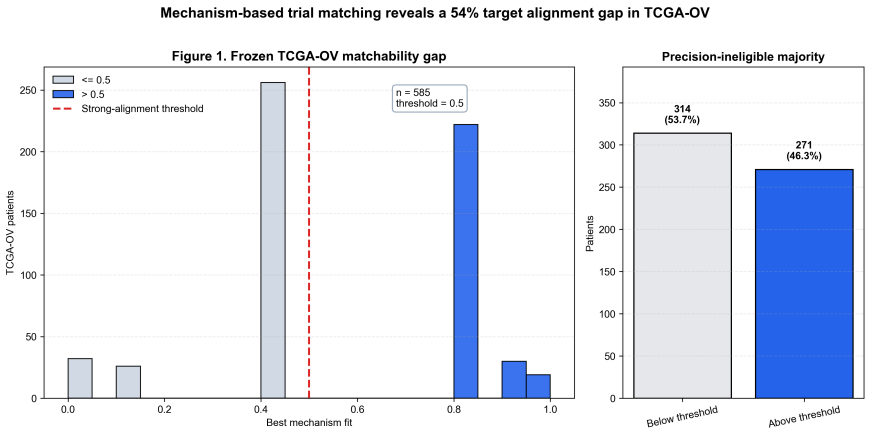

Figure 1: Figure 1. Frozen TCGA-OV matchability gap with locked strong-alignment threshold `best_fit > 0.5`.

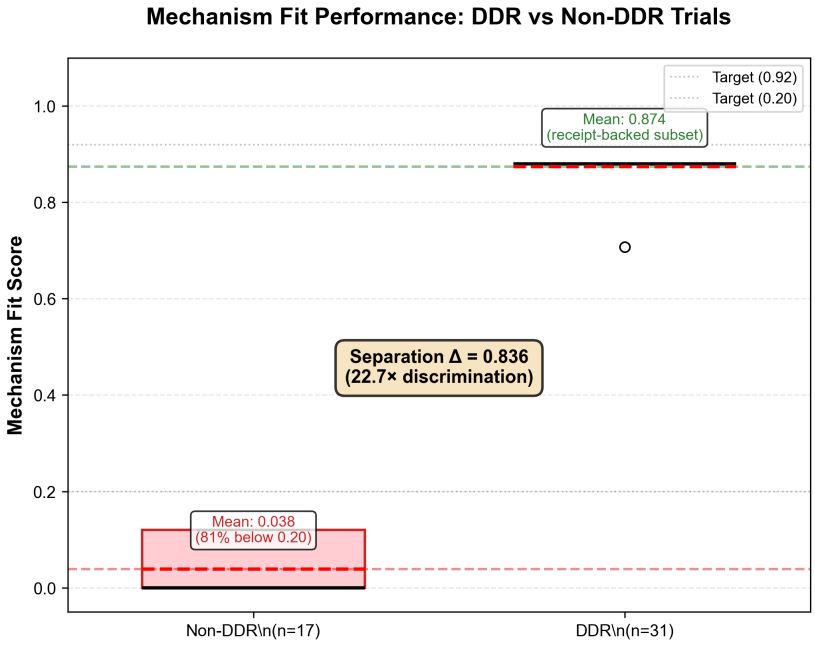

Figure 2: Figure 2. Discrimination of DDR-directed versus non-DDR trials on the curated 59-row menu.

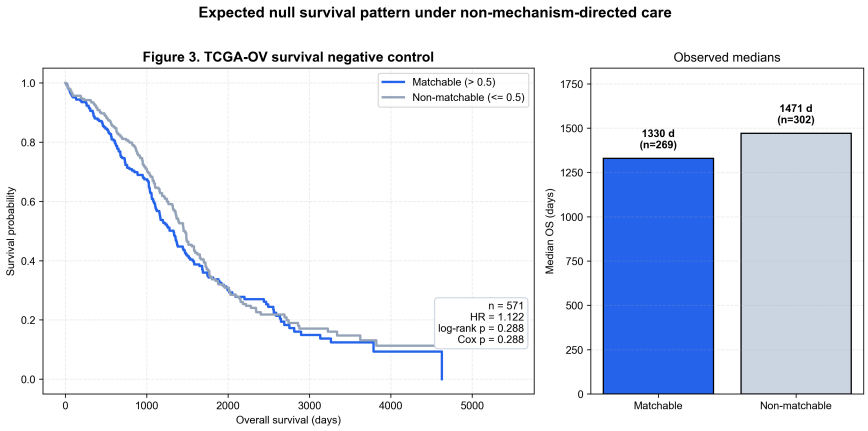

Figure 3: Figure 3. Kaplan-Meier negative control for matchable versus non-matchable groups on the frozen TCGA-OV branch.

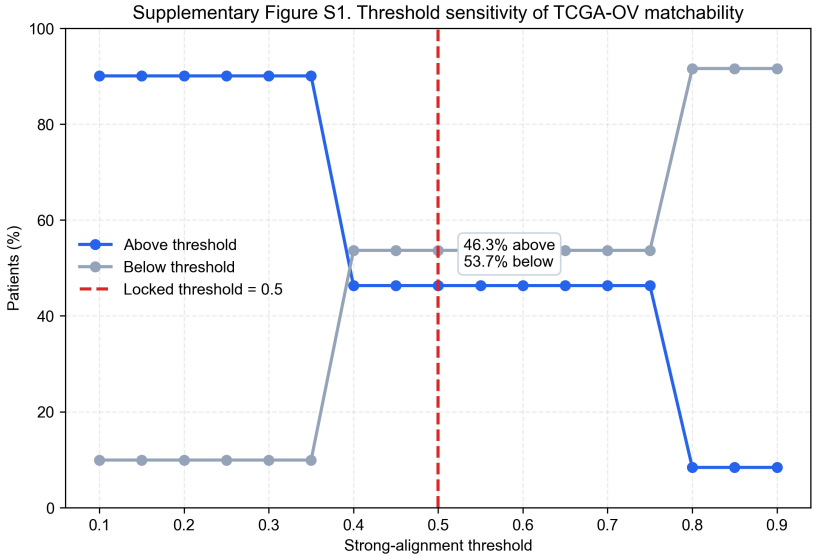

Figure 4: Supplementary Figure S1. Sensitivity of the prevalence split across strong-alignment thresholds.

Group	Count	Notes
Curated publication menu	59	Frozen 59-row ovarian cancer menu used for the manuscript analyses.
DDR-tagged binary rows	31	Rows with sanitized DDR axis value > 0.5.
Non-DDR binary rows	17	Rows retained in the 31 vs 17 discrimination split.
Zero-vector exclusions	11	Rows kept in the menu accounting but not used in the binary fit contrast because the 7D trial vector had zero norm.
Source: manual_intelligence_report	5	Curation source for rows included in the frozen manuscript menu.
Source: manual_keyword_matching	42	Curation source for rows included in the frozen manuscript menu.
Source: manual_zo_correction	2	Curation source for rows included in the frozen manuscript menu.
Source: openai_batch_tagging	10	Curation source for rows included in the frozen manuscript menu.