The Missing Mathematics of Biological Measurement

A theoretical perspective on causal displacement in fate-linked assays

Freedom Preetham*
Independent Researcher, CELL (Consortium for Equations of Life and Living-systems)
San Francisco, CA, USA | freedom@freedompreetham.org
* ORCID: 0009-0004-1278-8728

An ovarian cancer cell divides. One sister is sacrificed for single-cell RNA sequencing. The other receives olaparib. Days later, the second sister survives, and a shared lineage barcode connects that resistance to the transcriptome measured in the first.

The experiment is elegant because it recovers something ordinary single-cell sequencing destroys. It links a molecular description of the present to a functional outcome in the future. ReSisTrace used this design to identify pretreatment transcriptional features associated with resistance to olaparib, carboplatin, and natural killer cell exposure. The investigators deliberately limited the number of divisions because the molecular correspondence between related cells weakens as their trajectories separate.

The central insight of this essay is that fate-linked assays can silently change the biological unit being studied. Cell A supplies the molecular measurement, while cell B supplies the later outcome. Unless the two cells remain biologically equivalent for the intervention and time horizon under study, the experiment identifies a cross-cell relationship rather than a continuous same-cell trajectory. Cell B may inherit different proteins, organelles, chromatin states, or cell-cycle timing, and it may continue to diverge after division. The resulting association may capture a lineage-level propensity, yet it does not necessarily establish the future of the cell that was actually profiled. This distinction can change which biomarkers appear predictive, how resistance probabilities are estimated, and whether an added assay measures the biology of the profiled cell or merely the shared biology of related cells.

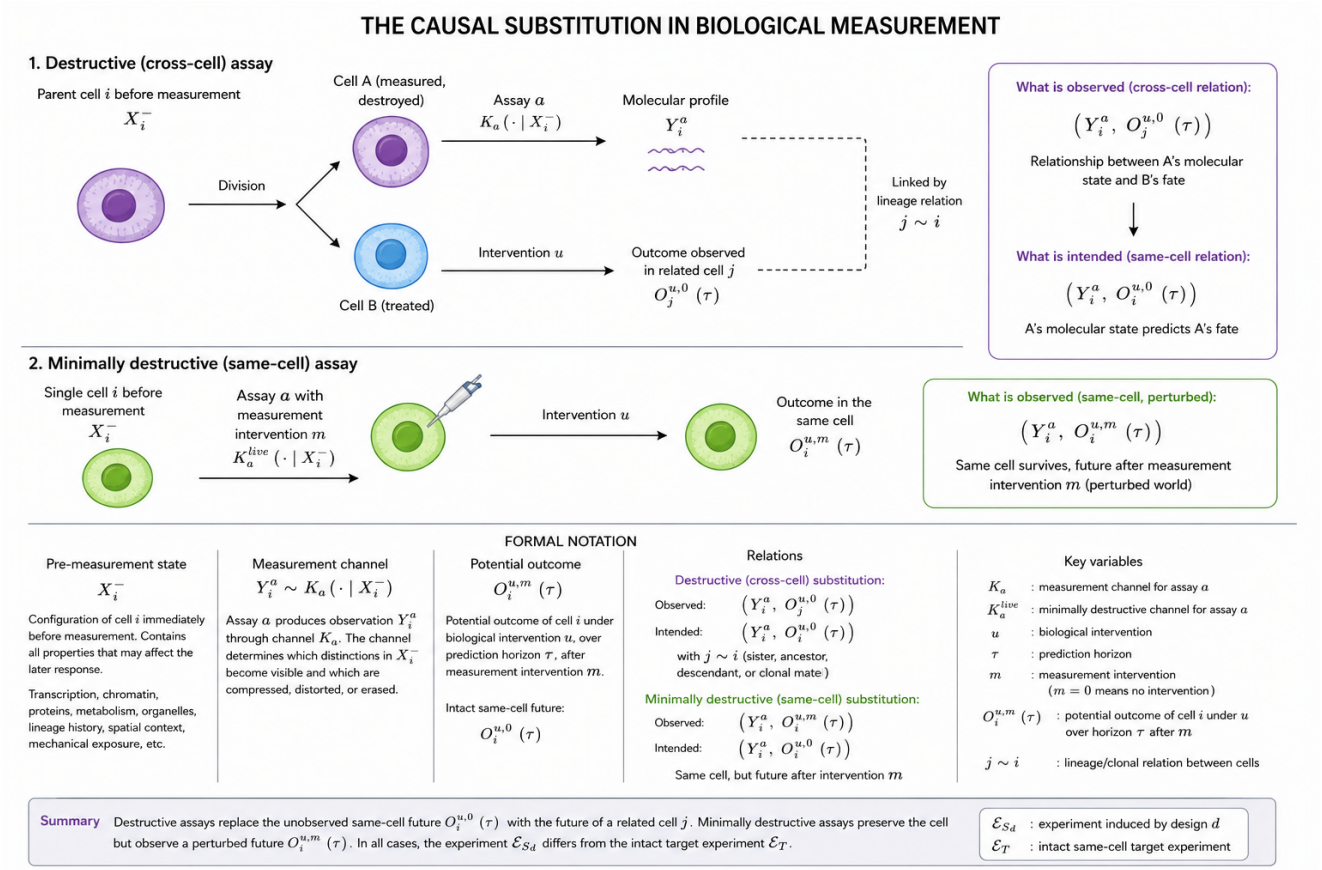
Perspective

Biology can measure the present so deeply that it destroys the future needed to interpret it.
Destructive assays profile one cell, then infer its fate from a sister, clone, descendant, or measurement-perturbed continuation. This substitution changes the measurand before noise, uncertainty, or model error are calculated. The observed relationship may become increasingly precise while the intended same-cell relationship remains weakly identified.

The missing mathematics must quantify this causal displacement. It must separate uncertainty introduced by the experimental design from the biological consequence of that displacement and from error introduced by the computational model. Without this separa-

tion, deeper sequencing, broader multiomics, and larger foundation models can achieve extraordinary accuracy within a proxy experiment while remaining uncertain about the biological future the study claims to explain.

The same substitution appears across modern biology. A clone mate supplies the phenotype of a cell that was lysed. A descendant supplies the fate of an ancestor that was sequenced. One tissue fragment is profiled while another receives treatment. A minimally destructive assay preserves the same cell, although its trajectory now unfolds after biopsy, labelling, repeated illumination, reporter insertion, or another physical intervention. In every case, the experiment performed and the biological relationship being claimed may belong to different statistical worlds.



A representational summary of the two causal substitutions formalized in the essay.

Formalizing the causal substitution

To make this substitution mathematically visible, we must separate the biological state that exists before measurement from the observation produced by the assay and from the future outcome the experiment is trying to explain. These objects are often treated as though they belong to one continuous trajectory. In a destructive experiment, they do not.

Begin with the cell before the assay touches it. Let

$$X_i^-$$

denote the configuration of cell i immediately before measurement. It contains every property that may affect the later response, including transcription, chromatin organization, protein modification, metabolism, organelle state, lineage history, spatial context, and mechanical exposure.

Assay a produces an observation through the measurement channel

$$Y_i^a \sim K_a\left(\cdot \mid X_i^-\right)$$

The channel K_a determines which distinctions in the pre-measurement state become visible and which are compressed, distorted, or erased.

Let

$$O_i^{u,m}(\tau)$$

denote the potential outcome of cell i under biological intervention u, over prediction horizon τ, after measurement intervention m. The intact same-cell future is

$$O_i^{u,0}(\tau)$$

A destructive assay gives us Y_i^a and removes $O_i^{u,0}(\tau)$ from observation. A sister-cell design substitutes the future of a related cell

$$O_j^{u,0}(\tau)$$

where j is connected to i through sisterhood, ancestry, or clonal membership.

The observed relation is

$$\left(Y_i^a, O_j^{u,0}(\tau)\right)$$

The intended same-cell relation is

$$\left(Y_i^a, O_i^{u,0}(\tau)\right)$$

A minimally destructive same-cell method creates a different substitution

$$(Y_i^a, O_i^{u,m}(\tau))$$

The biological unit survives, while its future now belongs to a measured and potentially perturbed cell.

Live-seq demonstrated that cytoplasmic RNA could be extracted through fluidic force microscopy while preserving cell viability, allowing a transcriptomic sample to be connected to later behavior in the same physical cell. That achievement opens a different experimental world. Molecular coverage, throughput, physical perturbation, and possible downstream effects remain part of the measurement problem. Live-seq removes cross-cell transport and introduces cross-intervention transport.

There is no untouched anchor. Biology has a network of incomplete experimental worlds. Each preserves a different portion of the present and supplies a different version of the future.

Experimental displacement is partially identified

Let

$$\mathcal{E}_{S_d}$$

denote the statistical experiment generated by design d. It includes the assay, proxy relation, measurement intervention, biological intervention, timing, and outcome actually observed.

Let

$$\mathcal{E}_T$$

denote the intact same-cell target experiment. This is the experiment that would connect response-relevant information in the pre-measurement state to the unperturbed future of that same cell.

Both experiments must be indexed by a restricted family of biological possibilities

$$\theta \in \Theta_{\mathcal{D},Q}$$

The domain $\mathcal{D}$ specifies the biological system, intervention, outcome, horizon, context, and scale. The family Q specifies the response mechanisms or state distinctions that the experiment is intended to resolve.

This restriction is unavoidable. A comparison over every conceivable molecular configuration and every possible future would produce a worst-case statement too broad to guide an experiment. Biological resolution exists relative to a scientific question.

Within that domain, one-sided Le Cam deficiency provides an ideal measure of experimental displacement

$$\delta_{\mathcal{D},Q}\left(\mathcal{E}_{S_d} \to \mathcal{E}_T\right) = \inf_{M} \sup_{\theta \in \Theta_{\mathcal{D},Q}} \mathrm{TV}\left(MP_\theta^{S_d}, P_\theta^T\right)$$

The stochastic kernel M searches for the best transformation capable of converting observations from the source experiment into observations from the target experiment. A small deficiency means that the source preserves the distinctions available in the target within the declared biological problem. A large deficiency means that the assay, proxy relation, or measurement intervention has erased distinctions required by the intended experiment. Le Cam's comparison theory was developed to compare statistical experiments through their probability laws and the decisions they can support.

The exact deficiency is generally unidentifiable because $\mathcal{E}_T$ is physically unavailable. It should never enter a protocol ranking as though it were an observed scalar.

The operational object is an identified interval

$$\mathcal{I}_d = \left[\underline{\delta}_d, \overline{\delta}_d\right]$$

Its width is

$$W_d = \overline{\delta}_d - \underline{\delta}_d$$

The upper bound represents the largest task-restricted displacement compatible with the observed experiments and declared assumptions. The width records how uncertain that displacement remains.

A design with an attractive lower bound and a wide interval has established little. It may approximate the target closely, or it may remain far from it. The experiment cannot yet distinguish those possibilities.

Sequencing depth can narrow uncertainty inside the sister-cell experiment while leaving the interval to the same-cell target almost unchanged. Molecular precision increases inside the substituted experiment. Same-cell identifiability does not automatically follow.

The limits of experimental anchors

Different experimental worlds can constrain one another because their errors differ. Sister-cell experiments carry cross-cell transport. Minimally destructive methods carry cross-intervention transport. Longitudinal imaging preserves continuity while observing a narrow molecular projection. Terminal multiomics gains molecular depth after the trajectory has ended.

Suppose $\mathcal{E}_A$ is a sister-cell experiment and $\mathcal{E}_B$ is a minimally destructive same-cell experiment. The compositional property of deficiency gives

$$\delta\left(\mathcal{E}_A \to \mathcal{E}_T\right) \leq \delta\left(\mathcal{E}_A \to \mathcal{E}_B\right) + \delta\left(\mathcal{E}_B \to \mathcal{E}_T\right)$$

The first edge may be constrained through overlapping populations, shared interventions, and paired outcomes. The second describes how strongly the proposed anchor changes the intact trajectory.

This relation becomes useful only when the intermediate experiment lies reasonably close to the target. Cytoplasmic extraction, reporter expression, repeated illumination, or another measurement effect may alter the fate under study. In that case

$$\delta\left(\mathcal{E}_B \to \mathcal{E}_T\right)$$

becomes large. The bound remains mathematically valid and experimentally uninformative. Preserving the same cell does not guarantee preservation of the same future.

Repeated siblings can reveal how response coherence decays after division. Biopsy-dose experiments can expose measurement-dependent changes in fate. Delayed interventions can identify when a proxy stops carrying information about the target horizon. Orthogonal outcomes may show strong transport for acute apoptosis and weak transport for long-term resistance.

Metrology advances by tightening uncertain edges between experimental worlds. No node acquires the status of ground truth merely because its intervention appears less destructive.

Experimental displacement, decision consequence, and model error

Experimental displacement matters through a declared biological decision. Predicting acute cell death, selecting a therapy, estimating lineage commitment, and forecasting durable resistance require different loss functions.

Because the target experiment is partially identified, its decision consequence must also remain partially identified. Let

$$\Delta_L(P)$$

denote the consequence of source-to-target displacement under loss L for an admissible data-generating process P.

The identified decision interval is

$$\mathcal{J}_{L,d} = \left[\underline{\Delta}_{L,d}, \overline{\Delta}_{L,d}\right]$$

where

$$\underline{\Delta}_{L,d} = \inf_{P \in \mathcal{G}_d(\alpha)} \Delta_L(P)$$

and

$$\overline{\Delta}_{L,d} = \sup_{P \in \mathcal{G}_d(\alpha)} \Delta_L(P)$$

Its width is

$$W_{L,d} = \overline{\Delta}_{L,d} - \underline{\Delta}_{L,d}$$

The family $\mathcal{G}_d(\alpha)$ contains the data-generating processes compatible with the observed experiments, calibration data, and prespecified structural hypotheses.

A wide displacement interval may yield a narrow decision interval when the unresolved differences do not affect the chosen action. A statistically modest shift may produce a severe consequence when it changes a rare resistant population or reverses a treatment decision. Experimental displacement and biological consequence remain distinct, while both inherit uncertainty from the unavailable target.

The fitted computational model introduces a third quantity

$$\varepsilon_M = \mathcal{R}_{S_d}(\widehat{g}) - \mathcal{R}^{\star}_{S_d}$$

This is the excess risk of the empirical predictor relative to the best attainable predictor within the experiment that was actually performed.

These quantities locate different failures. A large or weakly identified displacement interval implicates the experimental design. A consequential decision interval shows that the displacement matters for the biological question. Large model excess risk indicates that the algorithm failed to extract information already present in the source experiment.

A foundation model cannot compute its way out of a wide displacement interval. An assay developer cannot blame the model after the source experiment has erased the required distinction. A model developer cannot blame the assay while empirical risk remains far above attainable source risk.

Structural assumptions remain provisional

The displacement interval cannot be narrowed without assumptions about which mechanisms remain shared across experimental worlds. In cell biology, those mechanisms are rarely known well enough to function as fixed laws. The gene regulatory network may be the object under investigation. A graph inferred from the assay cannot become independent causal truth because it makes transport identifiable.

Let

$$\{\mathcal{G}_{d,q}(\alpha)\}_{q\in Q}$$

represent admissible model classes indexed by structural hypothesis q. One may assume that recent sisters remain exchangeable over a short horizon. Another may permit bounded lineage divergence. A third may posit a low-dimensional response state or a sparse family of effect modifiers.

A conclusion becomes persuasive when the displacement and decision intervals remain narrow across several experimentally credible hypotheses. When different hypotheses produce different answers, the analysis has exposed genuine epistemic instability.

New evidence may invalidate the current hypothesis family. The updated model class can expand

$$\mathcal{G}_t \subset \mathcal{G}_{t+1}$$

and the displacement interval can widen

$$W_{d,t+1} > W_{d,t}$$

Such widening represents discovery when the previous model family was too narrow. Experimental design should never reward only geometric interval contraction. That objective would avoid the experiments most capable of exposing missing mechanisms.

Batch experimental design under prior uncertainty

Biological experiments are slow. A resistance or differentiation protocol may require weeks before its outcome appears. The practical unit of adaptation is therefore a batch of complementary experiments run in parallel.

Let

$$H = (q, \psi)$$

represent a structural hypothesis q and its transport parameters ψ. Let $\mathcal{D}_t$ denote the observations available before batch t.

Expected information gain depends on the prior over H. This dependence becomes consequential in a partially identified causal space because the data may never overwhelm the initial structural commitments. The ranking of candidate experiments under expected information gain can be sensitive to prior perturbations, which motivates robust Bayesian design using ambiguity sets.

A single maximum-entropy prior does not solve the problem. Entropy depends on parameterization, and no prior can represent mechanisms absent from the hypothesis space. A stronger approach begins with a prespecified family of structurally diverse priors.

Let

$$\Pi_t$$

denote the family of posterior distributions induced by that prior set after observing $\mathcal{D}_t$. Let μ_t be a declared weighting measure over Π_t.

For each

$$\pi_t \in \Pi_t$$

the posterior predictive distribution of candidate batch $\mathcal{B}$ is

$$p_{\pi_t}\left(Y_{\mathcal{B}} \mid \mathcal{B}, \mathcal{D}_t\right) = \int p\left(Y_{\mathcal{B}} \mid H, \mathcal{B}\right) \pi_t(dH)$$

A hard minimax criterion can become paralyzing when the ambiguity set is broad. Almost every batch may be uninformative under at least one prior, leading the optimizer toward safe and uniformly mediocre experiments. The prior family must be broad enough to represent genuine structural disagreement and narrow enough to remain empirically credible. Its construction and sensitivity should be reported.

A lower-quantile robust information criterion offers a more practical compromise

$$U_{\mathrm{REIG}}^{(\beta)}\left(\mathcal{B} \mid \mathcal{D}_t\right) = \mathrm{Quantile}_{\beta, \pi_t \sim \mu_t}\left[I_{\pi_t}\left(H; Y_{\mathcal{B}} \mid \mathcal{B}, \mathcal{D}_t\right)\right]$$

The parameter β controls the required degree of robustness. A clinical program may choose a conservative lower quantile. Exploratory biology may accept greater prior sensitivity in exchange for experiments with higher upside. The full distribution of information gain across Π_t should accompany the selected batch.

This criterion still cannot discover a mechanism assigned zero support by every prior. Unknown biology remains outside the Bayesian loop until the experiment encounters something the model cannot explain.

Sentinel experiments are deliberately human-guided

Sentinel experiments test whether the current hypothesis family is adequate. They may involve an unusual cellular context, an orthogonal functional outcome, an extreme dose, an unexpected temporal window, or a perturbation selected because it violates the assumptions under which the current model is comfortable.

The current hypothesis family cannot derive its own sentinels reliably. By construction, the missing mechanism may lie outside every model and prior used by the optimizer. The sentinel set must therefore be assembled by investigators through biological judgment, failure analysis, adversarial review, prior anomalies, and knowledge of contexts that the current model treats poorly.

This is the point where human scientific judgment enters explicitly. It should be visible rather than disguised as an automated acquisition function.

Let

$$\mathcal{A}_{\text{sent}} \subseteq \mathcal{A}$$

denote the human-curated sentinel set. Let $C(\mathcal{B})$ be the total cost of batch $\mathcal{B}$, and let B_t be the available budget.

Require

$$C\left(\mathcal{B} \cap \mathcal{A}_{\text{sent}}\right) \geq \rho B_t$$

where

$$0 < \rho < 1$$

is the minimum fraction of experimental capacity reserved for model criticism.

A sentinel result outside the posterior predictive envelope triggers expansion or reconstruction of the hypothesis family. The displacement and decision intervals may widen afterward. That widening removes false certainty and belongs to scientific progress.

Decision-directed information gain

Information about structural hypotheses matters because it can alter a biological decision. The batch objective should therefore distinguish models whose differences lead to different actions.

For loss L, let $Z_L(H)$ denote the decision-equivalence class of hypothesis H. Two hypotheses belong to the same class when they induce the same Bayes-optimal biological action under the declared decision problem.

The decision-directed information gain of batch $\mathcal{B}$ under posterior π_t is

$$U_{\text{DIG}}\left(\mathcal{B} \mid \mathcal{D}_t, \pi_t\right) = I_{\pi_t}\left(Z_L(H); Y_{\mathcal{B}} \mid \mathcal{B}, \mathcal{D}_t\right)$$

Its lower-quantile robust form is

$$U^{(\beta)}_{\mathrm{RDIG}}\left(\mathcal{B} \mid \mathcal{D}_t\right) = \mathrm{Quantile}_{\beta,\pi_t\sim\mu_t}\left[U_{\mathrm{DIG}}\left(\mathcal{B} \mid \mathcal{D}_t, \pi_t\right)\right]$$

This rewards experiments expected to reveal which action is justified. It does not require uncertainty intervals to contract monotonically. A batch can widen the current interval and still carry high value when it exposes that the prior decision collapsed biologically distinct possibilities.

A feasible batch is selected from

$$\mathcal{B}^\star_t \in \mathrm{Pareto}\left\{ \left(U^{(\beta)}_{\mathrm{REIG}}\left(\mathcal{B} \mid \mathcal{D}_t\right), U^{(\beta)}_{\mathrm{RDIG}}\left(\mathcal{B} \mid \mathcal{D}_t\right), -C(\mathcal{B})\right) \;\middle|\; \begin{array}{l} \mathcal{B} \subseteq \mathcal{A} \\ C(\mathcal{B}) \le B_t \\ C\left(\mathcal{B} \cap \mathcal{A}_{\mathrm{sent}}\right) \ge \rho B_t \end{array} \right\}$$

No additional informational redundancy penalty is required. Joint mutual information already discounts experiments whose observations duplicate information contained elsewhere in the batch. Operational duplication, shared preparation burdens, and laboratory bottlenecks belong inside $C(\mathcal{B})$ or separate hard-capacity constraints.

The resulting batch contains experiments that distinguish transport models, experiments that discriminate among biological decisions, and human-selected sentinels included because the current hypothesis family may be wrong.

A harder standard for multimodal biology

Suppose RNA sequencing is expanded with chromatin accessibility, proteomics, morphology, or metabolism. The additional modality may improve prediction of the sister's response while leaving the interval to the sequenced cell's own future unchanged.

The model has improved inside the observed experiment. The causal claim has not moved closer to the intended experiment.

A modality should therefore report its incremental predictive value inside the source experiment and its effect on both uncertainty intervals

$$\Delta W_{b|a} = W_a - W_{a\oplus b}$$

$$\Delta W^L_{b|a} = W_{L,a} - W_{L,a\oplus b}$$

An expensive modality has failed to justify its same-cell claim when it improves source prediction while producing no robust reduction in experimental or decision uncertainty.

Ten thousand additional chromatin peaks may teach us more about the sister-cell dataset and almost nothing about whether the sequenced cell itself would have survived. Another million cells can estimate the proxy experiment with extraordinary precision while leaving

the target experiment poorly identified.

What a rigorous study should report

Candidate production designs should satisfy reservation requirements before entering final comparison

$$c_d \leq c_{\max}$$

$$p_d \leq p_{\max}$$

$$W_d \leq W_{\max}$$

$$\overline{\delta}_d \leq \delta_{\max}$$

The scientific program must set these values before comparing protocols. A developmental study may impose a severe perturbation limit. A clinical decision may demand a narrow decision interval. A discovery program may accept greater displacement in exchange for molecular breadth, although that compromise should remain visible.

A fate-linked study should state whose molecular state was measured, whose future was observed, which measurement world produced that future, and how tightly the observed experiment is known to approximate the intended one. It should report the experimental displacement interval, the downstream decision interval, and the model's excess risk as separate quantities.

The resulting claim might read as follows.

Within this cell system, under this intervention and over this horizon, the sister-cell experiment approximates the intact same-cell experiment with displacement contained in this interval. The associated decision consequence lies within this second interval. The added modality improves proxy prediction by this amount and changes both intervals by these amounts. The next batch was selected for structural and decision-directed information gain across a calibrated family of priors, with a fixed fraction of capacity reserved for human-selected experiments capable of falsifying the current transport model.

Molecular precision tells us how accurately an assay reads the world it creates. Causal precision depends on how tightly that world can be connected to the future the experiment erased.

Scope and mathematical provenance

ReSisTrace provides the biological example that motivates this essay. Its sister-cell design measures the molecular state of one cell and observes the later treatment outcome of a related cell. The mathematical framework presented here does not appear in the ReSisTrace study and was not derived from its experimental data.

This essay is a theoretical perspective that proposes a formal framework for measurement-induced causal displacement in dynamic biology. The equations combine established concepts from potential-outcome causal inference, causal transportability, Le Cam's comparison of statistical experiments, partial identification, decision theory, and Bayesian optimal experimental design. Their organization around displacement intervals, decision-consequence intervals, sentinel experiments, and multimodal assay design is proposed here. These equations define theoretical estimands and experimental-design criteria. They should not be interpreted as empirically validated biological laws or fitted models.

References

The references below document the biological examples and the established mathematical foundations used in this theoretical perspective.

1. Dai, J., Zheng, S., Falco, M. M., et al. (2024). Tracing back primed resistance in cancer via sister cells. *Nature Communications*, 15, 1158. https://doi.org/10.1038/s41467-024-45478-7

2. Chen, W., Guillaume-Gentil, O., Rainer, P. Y., et al. (2022). Live-seq enables temporal transcriptomic recording of single cells. *Nature*, 608, 733-740. https://doi.org/10.1038/s41586-022-05046-9

3. Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469), 322-331. https://doi.org/10.1198/016214504000001880

4. Pearl, J., & Bareinboim, E. (2014). External validity: From do-calculus to transportability across populations. *Statistical Science*, 29(4), 579-595. https://doi.org/10.1214/14-STS486

5. Le Cam, L. (1964). Sufficiency and approximate sufficiency. *The Annals of Mathematical Statistics*, 35(4), 1419-1455. https://doi.org/10.1214/aoms/1177700372

6. Torgersen, E. (1991). *Comparison of Statistical Experiments*. Cambridge University Press.

7. Blackwell, D. (1951). Comparison of experiments. In J. Neyman (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (pp. 93-102). University of California Press.

8. Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer.

9. Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4), 986-1005. https://doi.org/10.1214/aoms/1177728069

10. Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 10(3), 273-304. https://doi.org/10.1214/ss/1177009939

11. Foster, A., Jankowiak, M., Bingham, E., Horsfall, P., Teh, Y. W., Rainforth, T., & Goodman, N. (2020). Variational Bayesian optimal experimental design. In *Advances in Neural Information Processing Systems 33* (pp. 14036-14047).

12. Go, J., & Isaac, T. (2022). Robust expected information gain for optimal Bayesian experimental design using ambiguity sets. In *Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence*, Proceedings of Machine Learning Research, 180.

13. Liepe, J., Filippi, S., Komorowski, M., & Stumpf, M. P. H. (2013). Maximizing the information content of experiments in systems biology. *PLoS Computational Biology*, 9(1), e1002888. https://doi.org/10.1371/journal.pcbi.1002888

14. Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.