

Separating Biological and Choice Effects in Randomized Trials: Completing Causal Inference with β -Identification

Ogan Gurel

ogan.gurel@uta.edu

The University of Texas at Arlington

James Weinstein

Microsoft (United States)

Research Article

Keywords: preference-sensitive trials, causal estimands, instrumental variables, local average treatment effect (LATE), treatment crossover, nonadherence, patient choice, clinical trial design, clinical equipoise, shared decision making

Posted Date: August 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10011930/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Separating Biological and Choice Effects in Randomized Trials:

Completing Causal Inference with β -Identification

Ogan Gurel MD, The University of Texas at Arlington*

James N. Weinstein MS DO, Microsoft Research†

Keywords: preference-sensitive trials; causal estimands; instrumental variables; local average treatment effect (LATE); treatment crossover; nonadherence; patient choice; clinical trial design; clinical equipoise; shared decision making

* Corresponding author: Department of Bioengineering, 500 UTA Blvd, Arlington, TX 76010 ogan.gurel@uta.edu +1 (940) 290-5327

† Microsoft Research, 55 Hardscrabble Lane, Lyme NH 03768 jim.weinstein@microsoft.com +1 (603) 286-0140

- Total word count (main text) = 5,363
- The original SPORT study (ClinicalTrials.gov NCT00000410) was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (U01-AR45444-01A1) and the Office of Research on Women's Health, National Institutes of Health.

Abstract

Background: Randomized controlled trials in preference-sensitive care yield attenuated or indeterminate estimates when crossover and nonadherence are present. In these settings, outcomes decompose into biological effects (which we denote α) and preference-mediated choice effects (β), but analysis remains α -centric and omits β as an estimand component. Formally incorporating β into the standard randomized trial framework restores the full outcome structure. Secondary analyses of trials such as SPORT show that β is identifiable and can exert a causal, clinically meaningful effect on physical outcomes, necessitating its explicit inclusion.

Methods: We decompose observed outcome differences into a biological effect (α), the causal effect of randomization, and a preference-mediated component (β) arising from treatment choice. With randomized assignment as an instrumental variable and observed treatment receipt, α and β are identified within standard instrumental variable and local average treatment effect (LATE) frameworks. Using a transitional exemplar such as SPORT, we link the identification of β to observed preference-mediated effects, thereby extending conventional estimands within causal inference. We then specify joint estimation and interpretation of α and β and define classification by magnitude relative to the minimal clinically important difference.

Results: Crossover and nonadherence function as sources of causal identification rather than bias. Joint estimation of α and β separates attenuated effects into weak biological efficacy versus substantial preference-mediated influence. Across trials beyond SPORT, similar intention-to-treat results map to distinct regions, revealing underlying causal structures with corresponding differences in interpretation.

Conclusions: β -identification extends, rather than replaces, standard randomized trial analysis by treating the causal effect of treatment choice as an estimand component. α -centric inference remains a valid special case when preference-mediated effects are negligible, corresponding to the α -dominant region, with β -dominant and mixed settings defined by non-negligible preference effects. Expanding the estimand set to $\{\alpha, \beta\}$ preserves the logic of randomization while providing a more complete account of trial outcomes, constraining the domain of α -centric inference, separating biological generalizability from context dependence, and shifting interpretation, intervention, and evaluation to follow the joint structure of biology and choice.

Trial registration: Not applicable.

1. Introduction

This paper addresses a central problem in modern clinical trials: when biological differences narrow, outcomes turn on biology and patient choice together. Under those conditions, interpretation falters, because unquantified choice effects collapse back into biology and obscure what is driving outcome differences. We therefore present a framework for quantifying the preference-mediated effect (β) in clinical trials, with α denoting the biological treatment effect.

1.1 The Paradox of Rigor without Resolution

Randomized trials can remain equivocal despite rigorous design when crossover, nonadherence, and heterogeneous uptake are treated as noise rather than as measured structure. A trial may be well conducted yet still fail to yield clear clinical guidance. The Spine Patient Outcomes Research Trial (SPORT) for lumbar disc herniation exemplifies this paradox. Its protocols were rigorous, yet its results were judged inconclusive.^{1,2,3} The same pattern recurs when multiple biologically credible treatments coexist and explanatory mechanism no longer determines clinical decision, a tension identified by Schwartz and Lellouch but not resolved in inference.⁴ While nothing is wrong with biological efficacy, the paradox arises when inference stops there. When patient choice materially shapes outcomes, α alone becomes insufficient, and valid inference must include β to capture the full structure of outcomes.

1.2 What SPORT Revealed

McCormick's critique of SPORT identified crossover-driven attenuation toward the null and the resulting conflation of treatment assignment with treatment received, but remained α -centric and therefore attributed outcomes to biology alone. It captured the path of attenuation, not its consequence: why the signal weakened, but not what the equivocal findings structurally meant.³

When biological contrasts narrow, outcome differences may be governed less by biology alone than by how patients choose among credible options. That shift remains uninterpretable unless the design and data structure preserve the choice effect. Secondary analyses of SPORT made this visible by estimating the patient choice effect (β) on a common clinical scale and showing that physical outcomes were jointly shaped by biology and choice.^{5,6} When the net biological contrast $\Delta\alpha$ approaches zero, crossover and nonadherence no longer function simply as biases to suppress; under appropriate identification conditions, they become the variation through which β is identified and

quantified.^{5,6} This reframing restores interpretability by aligning the estimands with the outcome-generating process. Further detail is provided in Supplement S8.

1.3 Why Simple Randomized–Observational Comparison is Insufficient

Discrepancies between randomized and observational estimates have long been recognized: Chalmers *et al.* provided an early example in benign prostatic hypertrophy.⁷ There, divergence was visible but not decomposable; observational contrasts fold together choice, clinical judgment, and unmeasured confounding.⁸ SPORT made the divergence decomposable. Randomized assignment, combined with observed choice on a common outcome scale, permitted separation of preference-mediated effects (β) from biology (α).^{5,6} What earlier studies treated as bias thus became causally legible. Randomized–observational comparison alone is therefore insufficient; α and β must be causally identified and measured on a common outcome scale before the divergence can be read and interpreted.

1.4 Why β Now: Complexity, Decision Support, and Patient Agency

Three forces make this reframing timely. First, aging, multimorbidity, and polytherapy narrow net biological contrasts while amplifying trade-offs,^{9–11} shifting practice and the trials that inform it toward equipoise.¹² Second, decision-support systems built on biologically centered evidence remain blind to patient preference unless its effect on outcomes is properly identified.^{13,14} Third, contemporary emphasis on autonomy, shared decision making, and informed choice has raised requirements that patient values and agency should enter both care and research,^{15–17} requirements reinforced by policy and statute.^{18–21} Once choice materially shapes outcomes, β must be identified; otherwise trials may remain statistically precise yet causally uninterpretable.

1.5 From α -Centric Evidence to β -Identifying Inference

In preference-sensitive care, outcomes are jointly shaped by α and β , yet most trials remain α -centric.^{5,6} When the biological contrast is large ($|\Delta\alpha| \gg 0$), classical randomized inference remains sufficient and evidence-based medicine retains its usual force.^{22–24} When that contrast narrows toward zero ($\Delta\alpha \approx 0$), α -only interpretation no longer holds. Table 1 therefore classifies evidence along two orthogonal dimensions: what is identified in the evidence (α alone vs $\alpha + \beta$) and how that evidence is interpreted (β -blind vs β -aware). Quadrant I preserves the classical α -centric case, where conventional randomized inference remains sufficient. This paper concentrates on the β -identified domains (Quadrants III–IV), not to replace that case, but to extend interpretive scope where outcomes are jointly shaped by biology and choice.

Table 1. Regime-Conditional Evidence and Interpretive Structure

		INTERPRETATION	
		β -aware	β -blind
EVIDENCE	$\alpha + \beta$ (β -identified)	IV Rule-compliant integration β identifying trials - AI training mitigates α bias - Systems know their regime - Ranking only when α decides - Deferral + elicitation when β decides	III Human-legible, AI-inaccessible - Trials measure β correctly - Data contains β and $\Delta\beta$ on outcome scale - Clinicians can reason with β - Shared decision making is supported - AI still ranks by α and collapses β
	α only (α -centric)	II Transitional safety - Evidence still α-biased β uncertain or missing - System detects regime risk - Defers ranking under likely β-dominance - Explicit preference elicitation	I Classical randomized inference - Evidence reports α only - Systems rank under equipoise β effects are invisible β-blind ranking and false decisiveness - Conventional evidence-based medicine remains sufficient

Clinical trial evidence classified along two orthogonal dimensions: what is identified (biological effects alone vs biological and preference-mediated effects) and how estimates are interpreted (β -blind vs β -aware). Quadrants I–IV distinguish the status quo, transitional use, β -identified but misinterpreted evidence, and full alignment of measurement and interpretation under regime rules. In practice, because most clinical decision-support systems remain β -blind and β is not routinely identified, current use operates primarily in Quadrant III, where this structure is present but not incorporated into inference. Integration of β into decision-support and β -aware inference is addressed in related work.^{13,14}

1.6 Enabling β -Identifying Inference

Classical randomized inference treats deviations from randomization—nonadherence and crossover—as noise rather than signal. β -identifying inference, in contrast, distinguishes biological effects (α), understood as observer-independent biological efficacy, from observer-dependent preference-mediated choice effects (β), which arise at the time of

decision and mark a different causal driver.^{5,6} Accommodating preference in design does not by itself identify β ; β -identifying inference requires β to be prespecified and measured as a clinical outcome-scale estimand.^{5,6,25,26} Without that separation, outcomes carried by choice may still be misread as weak biology, and trials may remain equivocal even when choice governs them. We now describe the methodology for β -identifying inference.

2. Methodology

Observed outcomes in preference-sensitive trials arise from both biological effects and treatment choice. Accordingly, we define co-estimands (α , $\Delta\alpha$, β) on a common clinical scale and treat β as a causal component rather than residual variation, locating trials in ($|\Delta\alpha|$, β) space. Identification proceeds under randomized assignment, with treatment receipt providing the channel through which β is estimable. Interpretation, design, and power are then determined by this joint structure. The main text specifies the governing structure; the Supplement details implementation.

2.1 Axioms (A): Invariants of Preference-Sensitive Care

Inference rests on structural principles. In preference-sensitive settings, outcomes are determined by both biological effects and preference-mediated choice, and interpretation must reflect both. This introduces a boundary condition on measurement: preference can enter inference only if its elicitation does not itself act as an intervention on treatment uptake. If elicitation alters uptake, β is no longer interpretable as a causal component. Given these conditions, six axioms of β -identifying inference follow.

- **A1 – Dual Causation ($\alpha + \beta$):** Outcomes arise jointly from biological effects and patient choice.^{5,6}
- **A2 – Estimand–Data Concordance:** Under nonadherence, α must be causally identified (e.g., IV/LATE/CACE) and β defined on the same outcome scale.^{8,27}
- **A3 – MCID Discipline:** Interpretation is anchored to clinical materiality, not statistical significance alone.²⁸
- **A4 – Regime Rule ($|\Delta\alpha|$, β):** Interpretation and design follow the joint magnitude of $|\Delta\alpha|$ and β .
- **A5 – Identifiability Before Optimization:** α and β must be causally identified before estimation, design tuning, or decision making.^{8,29,30}
- **A6 – Non-Interference of Elicitation:** Preference measurement must not influence treatment uptake.^{15,31}

Most of these axioms will be familiar to trialists. Well established principles include A2 (Estimand–Data Concordance) and A3 (MCID Discipline) which are standard in trial analysis.^{5,6,28} A5 (Identifiability Before Optimization) is also conventional in form, but extends once β is treated as an estimand alongside α .^{8,29,30} In contrast, A1 (Dual Causation), while empirically based on the secondary analysis of SPORT,^{5,6} generalizes the structural model of outcomes by admitting β explicitly, and A4 (Regime Rule) introduces a classification absent from classical randomized inference, which operates implicitly within the α -dominant case.

A6 (Non-Interference of Elicitation) follows from the distinct status of β : unlike α , which is observer-independent and defined ex ante, β depends on observed choice at decision time and therefore requires that preference measurement not act as an intervention. This condition is partially grounded in prior work: research ethics emphasize non-interference with participant welfare and behavior,¹⁵ while studies of shared decision making show that elicitation can actively shape treatment choice.³¹

The novelty lies in the joint constraint: once β is admitted as an outcome-scale estimand, identifiability, design, and interpretation are governed simultaneously by biological contrast and enacted choice. Formal statements and validity constraints (D1–D3) are provided in Supplement S1.

2.2 Canonical α and β Estimands

The framework rests on causal estimands defined on a common clinical outcome scale: the biological treatment effect (α), the net biological contrast between options ($\Delta\alpha$), the observational as-treated contrast (ΔOBS), and the preference-mediated effect (β), where $\beta = \Delta\text{OBS} - \Delta\alpha_{\text{causal}}$. A common scale places biological contrast and preference effect in one interpretive domain, allowing direct comparison against the same clinical threshold. Under nonadherence, α is identified through randomized assignment as an instrument (LATE/CACE), whereas β also requires observed treatment choice and adequate IV validity and strength.^{8,27} β further requires that preference elicitation remain measurement rather than intervention.^{15,31} Table 2 sets out these co-estimands; Supplement S2 provides detailed definitions, estimators, and diagnostics, and Supplement S3 the formal identification conditions.

Table 2. Canonical Co-Estimands:
the Biological Effect (α), the Net Contrast ($\Delta\alpha$), and the Choice Effect (β)

		Meaning (clinical scale)
Biological (Treatment) Estimands	α	Biological treatment effect (assignment or causal effect under nonadherence)
	$\Delta\alpha$	Net biological contrast between treatment options
	ΔOBS	Observed (as-treated) outcome difference
Preference Estimands	β	Preference-mediated effect $\Delta\text{OBS} - \Delta\alpha_{\text{causal}}$

Canonical estimands are reported on a common clinical outcome scale, enabling direct comparison of biological (α) and preference-mediated (β) effects. The biological effect (α) is estimated as Intention-to-Treat (ITT) and, under nonadherence, as LATE/CACE using instrumental variables;⁸ the net contrast ($\Delta\alpha$) is the difference between options on this scale; and β is defined as the residual between observational and causally identified contrasts. All estimands are reported on a common outcome scale for direct comparison of α , $\Delta\alpha$, and β . Interpretation is anchored to a prespecified Minimal Clinically Important Difference (MCID) (A_3).²⁸

2.3 Rules (R): Pre-specified Rules of Interpretation and Action

Interpretation and design are governed by prespecified rules that bind ($|\Delta\alpha|$, β) to inference, design, and action. These rules fix the threshold at which biological effects remain decisive, preference becomes consequential, and trials must be designed, powered, and interpreted differently across regimes. The five rules that follow operationalize β -identifying inference.

- **R1 – Interpretation Rule:** Interpretation follows the joint magnitude of $|\Delta\alpha|$ and β relative to clinical materiality.²⁸
- **R2 – Regime Rule:** Design is routed by regime to make causal structure identifiable.
- **R3 – Power Rule:** Trials must be powered for both α and β , not α alone.³²
- **R4 – Measurement Rule:** Preferences must be measured without affecting treatment uptake.^{15,31}
- **R5 – Complexity Drift Rule:** As complexity increases, $|\Delta\alpha|$ narrows and β becomes more influential.^{9,11}

Of the five rules, R1 and R4 rest on established principles of clinical interpretation and measurement;^{15,28,31} R3 and R5 extend these principles to joint estimands and regime dynamics;^{9,11,32} and R2 is a novel rule introduced by β -identifying inference, routing design by causal regime. Full statements of these rules are provided in Supplement S4.

2.4 Implementation (Minimal Methods Kit)

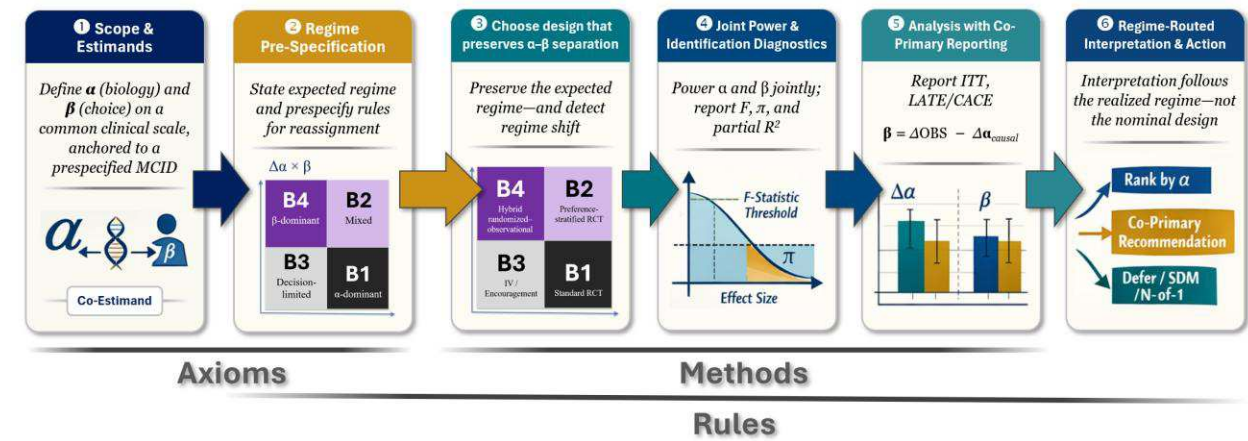
Implementation combines estimation of α and β with diagnostics of identification, uncertainty, and sensitivity. Under nonadherence, randomized assignment identifies α (LATE/CACE). Design may therefore be routed prospectively when mixed or β -dominant structure is plausible, but final regime assignment depends on identified estimates, prespecified thresholds, and diagnostics. The methods that follow operationalize that sequence.

- **M1 – Identify α :** ITT + LATE/CACE (random assignment as IV).^{8,27,32}
- **M2 – Estimate β (common scale):** $\beta = \Delta\text{OBS} - \Delta\alpha_{\text{causal}}$ (same outcome scale).^{5,6}
- **M3 – Is β Inference Permitted?:** Valid β inference requires a credible instrument (IV validity) and adequate first-stage strength; if either is doubtful, inference is decision-limited.^{8,33,34}
- **M4 – How Precise is the Inference?:** Quantify uncertainty and robustness using standard errors/resampling and prespecified sensitivity analyses.^{32,35}
- **M5 – Conditions for Valid β Inference:** Measure preference without affecting treatment uptake (elicitation must not function as an intervention).^{15,31}
- **M6 – Joint Power:** Joint α/β power, reflecting instrument strength.^{8,32}

In practice, β -identifying inference can be implemented in four steps: prespecify α and β on a common outcome scale; preserve observable treatment uptake rather than enforce adherence; estimate α under nonadherence using standard IV methods; and report as-treated and causal contrasts so that β can be computed directly. Full implementation details are provided in Supplement S5–6.

2.5 Putting it all together: Six stages from design to interpretation

The Axioms, Rules, and Methods introduced above are distinct but coordinated parts of the framework. Figure 1 shows how they are distributed across the life of a β -identifying trial. Axioms define what must structurally hold, Rules govern design and interpretation, and Methods make the resulting causal structure empirically legible. The six stages do not add a new theoretical layer; they show how A–R–M are coordinated in practice.

Figure 1. Six stages of a β -identifying trial

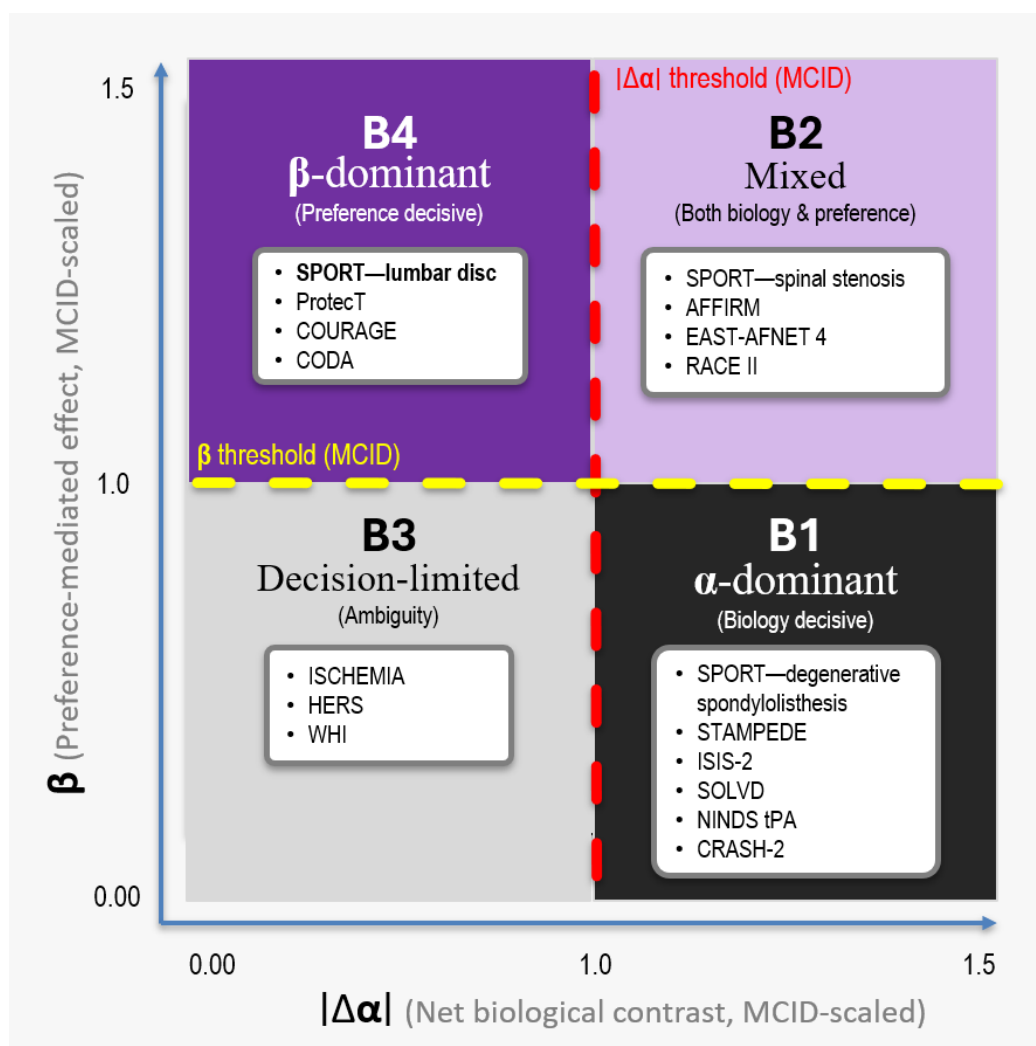
Operational workflow from design to interpretation. **Stage 1**, Scope and estimands: define α and β on a common clinical scale and anchor interpretation to a prespecified MCID. **Stage 2**, Regime pre-specification: state the expected regime and prespecify rules for regime shifts or uncertainty. **Stage 3**, Design selection: choose a design that preserves identifiability of the expected structure (e.g., standard randomized, preference-stratified, encouragement/IV, or hybrid randomized-observational). **Stage 4**, Joint power and identification diagnostics: power α and β jointly and prespecify diagnostics such as instrument strength and uptake variation. **Stage 5**, Analysis with co-primary reporting: report ITT, LATE/CACE, ΔOBS , and β on a common scale with uncertainty and sensitivity diagnostics. **Stage 6**, Regime-routed interpretation and action: interpret the realized trial through its assigned regime rather than its nominal design, and route action accordingly. **Axioms** anchor Stages 1–2; **Rules** govern Stages 2–6; **Methods** operationalize Stages 3–5.

3. Results

Here we apply the framework by locating trials in $(|\Delta\alpha|, \beta)$ space. This yields regime classification, exemplar placement, and rule-guided design. Regime assignment depends on the identified $(|\Delta\alpha|, \beta)$ structure and is therefore an inferential conclusion rather than a design label.

3.1 Regime map, classification & implications

Regime map. Regime classification follows by comparing the net biological contrast $(|\Delta\alpha|)$ and the preference-mediated effect (β) with prespecified MCID thresholds,²⁸ yielding four regimes: α -dominant (B1), mixed (B2), decision-limited (B3), and β -dominant (B4). Diagnostics belong with the estimates. Figure 2 displays this regime space and situates within illustrative trial exemplars, including SPORT (lumbar disc herniation) as an empirically grounded β -dominant anchor.^{1,2}

Figure 2. Regime Map ($|\Delta\alpha|$, β) Specifying Locus of Clinical Trials

Axes show the net biological contrast ($|\Delta\alpha|$) and the preference-mediated effect (β), each scaled to the minimal clinically important difference (MCID); dashed lines mark MCID thresholds on both axes.²⁸ The four regimes—B1 α -dominant, B2 mixed, B3 decision-limited, and B4 β -dominant—route interpretation, design, and clinical action by rule rather than convention. B1 supports α -first interpretation and standard randomized inference; B2 requires co-primary α/β interpretation and preference-accommodating designs; formally, B3 is the region in which both $|\Delta\alpha|$ and β remain below clinical decisiveness, but unlike the other regimes it also serves as the conservative assignment category when uncertainty or weak identification prevents confident placement in B1, B2, or B4; and B4 requires β -first interpretation and structured shared decision making. SPORT—lumbar disc (bolded)—is shown as the empirically grounded anchor.^{1,2} The remaining trial labels are illustrative; corresponding trial references are provided in Section 3.2; See Table 3 for the regime rulebook.

α -dominant. In B1, biology crosses clinical materiality while preference does not. Interpretation remains α -first, standard randomized design suffices, and action retains its classical form: recommend the biologically superior treatment. β may still be present, but remains descriptive because it does not govern the endpoint.

Mixed. In B2, both $|\Delta\alpha|$ and β cross clinical materiality: biology alone no longer settles interpretation, because choice enters as a co-primary determinant. Design must therefore preserve both—typically through preference-stratified randomization or partially randomized preference designs with joint α/β power—because collapsing the pathway back into α alone would erase the governing distinction. Recommendation shifts accordingly: it can no longer remain unconditional, but must proceed through shared decision making rather than β -blind ranking.

Decision-limited. Formally, B3 is the region in which both $|\Delta\alpha|$ and β remain below clinical decisiveness. Unlike the other regimes, it also functions as the conservative default when uncertainty or weak identification prevents confident assignment elsewhere. The consequence is diagnostic restraint: instrument validity must be strengthened, neutral elicitation improved, or power reallocated under M6 (Joint Power) before strong recommendation becomes legitimate. Until then, the study remains evidence-generating rather than decision-settling.

β -dominant. In B4, biological contrast falls below clinical decisiveness while preference-mediated structure crosses it. Interpretation therefore shifts from α -first to β -first. Once biology no longer settles the endpoint, design must turn with it: hybrid β -identifying or randomized–observational structures become necessary, shared decision making ceases to be optional, and, where within-person contrast is observable, N-of-1 inference becomes legible. In this regime, outcome no longer turns on treatment assignment alone, but on how choice is elicited, aligned, and enacted.

Complexity drift. As clinical complexity rises—through aging, multimorbidity, and polytherapy—biological contrasts narrow while preference effects grow.^{9–11} Thus clinical settings, and the trials built to study them, can move from α -dominant to mixed or β -dominant regimes (see Supplement S4 and Figure S1a). The shift is visible in anticoagulation for atrial fibrillation, where stroke-prevention decisions turn increasingly on bleeding risk, frailty, and adherence;³⁶ in coronary revascularization for complex disease, where symptom benefit must be weighed against procedural burden and patient goals;³⁶ and in localized prostate cancer, where clinically credible options yield preference-sensitive outcomes.^{38,39} Thus, as populations age and complexity accumulates, α -dominant findings from selective trial populations become more circumscribed and less transportable to complex real-world settings.^{4,40–42} In that interval, β -identifying designs preserve interpretability by separating biological from choice-driven structure before α -centric inference conflates them.^{5,6}

3.2 Placement of clinical trial exemplars

Figure 2 places selected trials in $(|\Delta\alpha|, \beta)$ space so that landmark studies can be read by regime rather than by conventional trial labels alone. These placements are interpretive and shown relative to MCID-based regime thresholds, not as exact standardized coordinates.

SPORT as anchor exemplar. SPORT (lumbar disc herniation) sits in the β -dominant regime (B4) as the clearest exemplar.^{1,2} It is read not simply as a crossover problem but as a hybrid randomized–observational structure in which substantial crossover and informed choice render β visible under near biological equipoise. SPORT is unique because its B4 placement is not merely interpretive but empirically grounded: the SPORT design and subsequent analyses permit explicit separation of biological treatment effects (α) from the causal contribution of enacted patient choice (β).^{5,6}

Relative to the lumbar disc herniation anchor, the other SPORT conditions are structured estimates. SPORT (spinal stenosis) falls in the mixed regime (B2): surgery improved outcomes over nonoperative care, yet crossover remained high and outcomes still turn on symptom burden, recovery demands, and willingness to undergo decompression.^{43,44} SPORT (degenerative spondylolisthesis) shifts toward α -dominant (B1) because surgery produced larger and more durable gains in pain and function than nonoperative care ($|\Delta\alpha| \gg 0$).^{45,46}

Other clinical trial exemplars. Outside SPORT, these placements are structured estimates rather than direct regime assignments based on a quantified β . ProtecT sits in the β -dominant regime (B4) because prostate-cancer mortality differed little at 10 years even as monitoring, surgery, and radiotherapy entailed markedly different quality-of-life burdens and treatment trade-offs.^{38,39} CODA also sits in B4 because antibiotics were noninferior to appendectomy in the short term, yet later appendectomy risk, appendicolith status, and differing complications left outcomes contingent on patient-valued trade-offs.⁴⁷ COURAGE lies there more cautiously: PCI added to optimal medical therapy did not reduce death or myocardial infarction relative to optimal medical therapy alone, leaving symptom relief, procedural burden, and patient goals to govern action.⁴⁸

The mixed (B2) cluster marks trials in which biological treatment differences are clinically material, but patient choices also enter into the decision. AFFIRM, EAST-AFNET 4, and RACE II illustrate this structure in different ways: atrial fibrillation management still depends on biological effects on rhythm control, symptom burden, and downstream cardiovascular risk, yet outcomes also bear the burdens of hospitalization, adverse effects, treatment timing, and what patients can realistically sustain.^{49–51} ISCHEMIA sits in the decision-limited regime (B3) because the evidence does not bind $|\Delta\alpha|$ strongly enough for decisive ranking, and β is not identified strongly enough to govern interpretation.³⁷ HERS and WHI also fall in B3: each

overturned an assumed biological superiority without identifying an alternative action-governing structure.^{52,53}

Design consequences. The β -dominant and mixed exemplars in Figure 2 identify settings in which future trials may need embedded β -identifying substudies or β -sensitive modules to preserve α - β separation. Otherwise, α -centric reporting may capture the biological pathway while leaving the action-governing structure only partly visible. Supplement S11 (Table S7) extends this logic to additional β -identification-ready clinical settings, several overlapping with the landmark trials shown here.

3.3 Regime assignment and rule-guided design

Once trials are placed in $(|\Delta\alpha|, \beta)$ space, interpretation and design become rule-governed. Table 3 specifies what may be concluded, what design remains sufficient, and what action or decision support remains justified. Table 4 then extends that assignment into design archetypes by linking each regime to its inferential aim, required design features, and methods of implementation.

Regime assignment. Regime assignment converts map placement into rule-governed interpretation. Once $|\Delta\alpha|$ and β are positioned relative to MCID, Table 3 supplies the operational rulebook: α -dominant when biological contrast is clinically material but preference remains weak, mixed when both are material, β -dominant when preference is material under near biological equipoise, and decision-limited when neither is decisive or uncertainty prevents confident assignment. Assignment is therefore not narrative judgment but inferential conclusion. When estimates lie clearly on one side of MCID, assignment proceeds directly; otherwise it defaults conservatively to the decision-limited regime until the evidence is more decisive. Supplement S7 formalizes that logic under uncertainty. Table 3 serves, therefore, not as a descriptive taxonomy but as an action rule: regime assignment determines how the trial may be interpreted, whether the design is sufficient for that interpretation, and what recommendation or decision support remains justified.

Table 3. Regime Rulebook ($\Delta\alpha \times \beta$): Interpretation, Design, and Action

	$ \Delta\alpha \approx 0$ (Biological equipoise)	$ \Delta\alpha \gg \text{MCID}$ (Clear biological superiority)
β large / strong (Preference materially affects outcomes)	B4 β-Dominant Interpretation β-first - Biology equivocal Design requirement β-identifying or hybrid randomized–observational designs - Explicit preference elicitation - N-of-1-compatible where feasible Clinical action - Preference-first decisions - Structured SDM required - AI must defer unless β is explicitly elicited and incorporated	B2 Mixed Interpretation α and β are co-primary - Biology matters, but choice modifies outcomes. Design requirement - Preference-stratified or PRPP-type designs - Joint α/β power - Preserve separation of biology and choice. Clinical action - Conditional recommendations - Structured SDM - No unconditional α-only ranking - AI must defer unless β is explicitly elicited and incorporated
β small / weak (Minimal preference signal)	B3 Decision-Limited Interpretation - Indeterminate - Neither biology nor preference is yet clearly decisive Design requirement - Strengthen instrument - Improve elicitation, or re-power - Treat apparent nulls cautiously Clinical action - No strong recommendation - Evidence-generating only	B1 α-Dominant Interpretation α-first; - Outcomes driven primarily by biology Design requirement - Standard randomized trial sufficient β may be reported descriptively Clinical action - Recommend biologically superior treatment - Conventional α-based decision support remains appropriate - Grounding for evidence-based medicine

Classifies trial settings by the joint magnitude of the net biological contrast ($|\Delta\alpha|$) and the preference-mediated effect (β), each evaluated relative to MCID, and specifies the interpretation, design consequence, and clinical action that follow from the assigned regime. The table is the operational rulebook of the framework: once a trial is placed in ($|\Delta\alpha|$, β) space, the corresponding cell governs what conclusions and recommendations remain justified. Horizontal axis: net biological contrast, estimated causally and read against MCID. Vertical axis: preference-mediated contribution to outcomes, estimated on the same clinical scale and read against MCID. When uncertainty obscures threshold crossing, assignment defaults conservatively to B3; see Supplement S7.

Rule-guided design archetypes. Once regimes are assigned, prespecified rules determine interpretation and design. Design is not interchangeable across regimes: what suffices in one may fail in another as the joint structure of $|\Delta\alpha|$ and β changes. Table 4 links each regime to its inferential aim, required design features, and methods of implementation, carrying regime assignment into the architecture of inference.

Table 4. Rule-Guided Trial Design Archetypes Across Regimes

Regime	Primary aim	β -identifying archetypes	Rules → Methods
B1 α -dominant	Confirm biological efficacy (α); β secondary	Conventional randomized controlled trials; ^{22–24} pragmatic or registry-embedded randomized trials ^{40–42,55–58} with secondary reporting of choice effects ^{25,26}	R1–R3 → M1–M2, M5
B2 mixed	Co-primary inference on α and β	Preference-stratified randomized trials (comprehensive cohort designs) ^{25,26,54} randomized encouragement with symmetric preference elicitation ^{8,27,31}	R1–R4 → M1–M2, M5–M6
B3 Decision-limited	Diagnose identifiability and adapt design	Instrument-strengthening ^{8,33,34} or preference-adaptive designs (e.g., enhanced encouragement, improved elicitation, re-estimation of power) ^{25,26,31,54}	R2–R3 → M4–M6
B4 β -dominant	Preference-first inference; estimate β	Hybrid randomized–observational designs (SPORT-class); aggregated N-of-1 trials; ⁵⁹ stepped-wedge; ⁶⁰ or sequential preference-eliciting designs	R1, R4 → M1–M2, M5, M6

The table maps empirically defined regimes (B1–B4) to their inferential aim, appropriate design archetypes, and the prespecified rules (R1–R5) and methods (M1–M6) required for valid implementation. Designs are not interchangeable across regimes: each is appropriate only where its identification assumptions hold. In β -identifying designs, preference effects are rendered empirically legible through the joint use of randomized assignment and structured observation of treatment choice on a common outcome scale, rather than inferred *post hoc* from crossover. Design families shown are grounded in classic explanatory–pragmatic distinctions and pragmatic “fit-for-purpose” guidance (B1),^{4,40–42} preference-trial and partially randomized preference designs (B2),^{25,26,54} instrument-based encouragement/identification logic under nonadherence and compliance heterogeneity (B3),^{8,27,33,34} and SPORT-class hybrid randomized–observational exemplars supporting β -quantification (B4).^{1,2,5,6}

3.4 β -Identifying Trial Designs

Canonical β -Decomposition. In preference-sensitive settings, β is quantifiable only when the observed treatment-received contrast (ΔOBS) and the causally identified biological contrast ($\Delta\alpha_{\text{causal}}$) are both available on the same clinical outcome scale. Under those conditions, β is defined by the canonical decomposition:^{5,6}

$$\beta = \Delta\text{OBS} - \Delta\alpha_{\text{causal}}.$$

The relation is simple, but its consequence is not: without aligned observational and causal contrasts, β cannot be read. In this decomposition, $\Delta\alpha_{\text{causal}}$ is not observational but causally identified under nonadherence through randomized assignment as an instrument (LATE/CACE); β is therefore defined only after the biological contrast has been separated from the observed treatment-received contrast.^{8,27}

Preference Accommodation Without β -Quantification. Preference accommodation is not yet β -quantification. In a patient-preference trial or comprehensive cohort design, patients with strong preferences choose treatment and only those without strong preferences are randomized; the preference arms therefore remain observational and do not separate preference from confounding.^{25,26,54} ProtecT shows the limit. Men who rejected randomization entered a treatment-choice cohort and were followed identically, but treatment preferences and treatments received were strongly associated with prostate-cancer risk features.^{38,39} The design therefore did not yield a decomposable pair of aligned contrasts— ΔOBS and $\Delta\alpha_{\text{causal}}$ —from which β could be computed. The missing element is not preference, but decomposability. Without both the observational and causal contrasts, the trial remains preference-aware rather than β -identifying. Hence β cannot be quantified.

Design Requirements for β -Identifying Trials. Once β is defined by decomposition, β -identifying design follows. Crossover and treatment switching are not merely bias to suppress; once α is identified, they can supply the variation through which β is quantified.^{5,6,8,27} A trial is β -identifying only if investigators (1) define α and β on a common outcome scale, (2) preserve observable treatment uptake, (3) estimate α under nonadherence using IV methods, and (4) report ΔOBS alongside $\Delta\alpha_{\text{causal}}$ so that β can be computed directly. In the transitional SPORT summary, both pathways retain favorable biological effects, the net biological contrast remains small ($\Delta\alpha \approx 0.65$), and β remains clinically meaningful ($\approx 6\text{--}8$ points) only because the observational and causal contrasts are both available in aligned form.^{5,6} Box 1 states the minimum operational requirements for a β -identifying randomized controlled trial.

Box 1 Minimal operational requirements for a β -identifying RCT

A standard randomized controlled trial satisfies the six-stage workflow (Figure 1) with the following minimal requirements. These do not introduce a new trial paradigm but ensure that the quantities necessary for β -identification are estimable and expressed on a common clinical scale.

Stages 1–2 (Scope, estimands, and regime specification)

- ✓ **Co-definition of α and β on a common clinical scale.** The biological contrast (α) and the preference-mediated contrast (β) are defined on an interpretable outcome scale and anchored to a prespecified MCID.
- ✓ **Prespecification of the expected regime and interpretation rules.** The anticipated structure (α -dominant, β -dominant, mixed, or decision-limited) and rules for regime assignment under uncertainty are specified a priori.

Stages 3–5 (Design, identification diagnostics, and estimation)

- ✓ **Preservation of randomization and ITT as the causal anchor.** Randomized assignment serves as the instrument identifying α . No modification of allocation, blinding, or primary endpoints is required.
- ✓ **Non-suppression of crossover and nonadherence.** Treatment uptake and switching are retained as observables. Enforcement beyond ethical or safety constraints is neither required for identification nor inferentially neutral to estimation.
- ✓ **Causal identification of α under incomplete adherence.** When adherence is incomplete, α is identified as LATE/CACE using randomized assignment, with prespecified diagnostics (instrument strength, uptake heterogeneity) and standard assumptions.
- ✓ **Co-primary reporting on a unified outcome scale.** Report ITT ($\Delta\alpha$), LATE/CACE ($\Delta\alpha_{\text{causal}}$), and the observed treatment-received contrast (ΔOBS) on the same clinical scale, with uncertainty and sensitivity diagnostics.
- ✓ **Mechanical definition of β .** The preference-mediated effect is defined as:

$$\beta = \Delta\text{OBS} - \Delta\alpha_{\text{causal}}.$$

β is inferred from outcomes and does not require direct measurement of preferences.

Stage 6 (Interpretation and action)

- ✓ **Regime-routed interpretation.** Interpretation follows the realized regime rather than the nominal trial design, with α and β evaluated relative to the MCID.
- ✓ **Explicit handling of decision-limited ambiguity.** When uncertainty spans clinically relevant thresholds, conclusions are classified as decision-limited rather than interpreted as null effects.

Implementation Note: Consistent with Stages 3–5, no novel estimation methods are required. Standard ITT/IV estimators identify α , standard as-treated contrasts estimate ΔOBS , and β is computed algebraically on a common scale. Under this structure, crossover is transformed from a perceived threat to validity into an identifiable component of the causal decomposition, with an inherent trade-off between biological precision and preference identification. Further details are provided in Supplement S1–S6 with formal regime assignment under uncertainty, including treatment of confidence intervals crossing MCID thresholds, specified in Supplement S7. The SPORT trial and its secondary analysis as a transitional exemplar is summarized in Supplement S9.

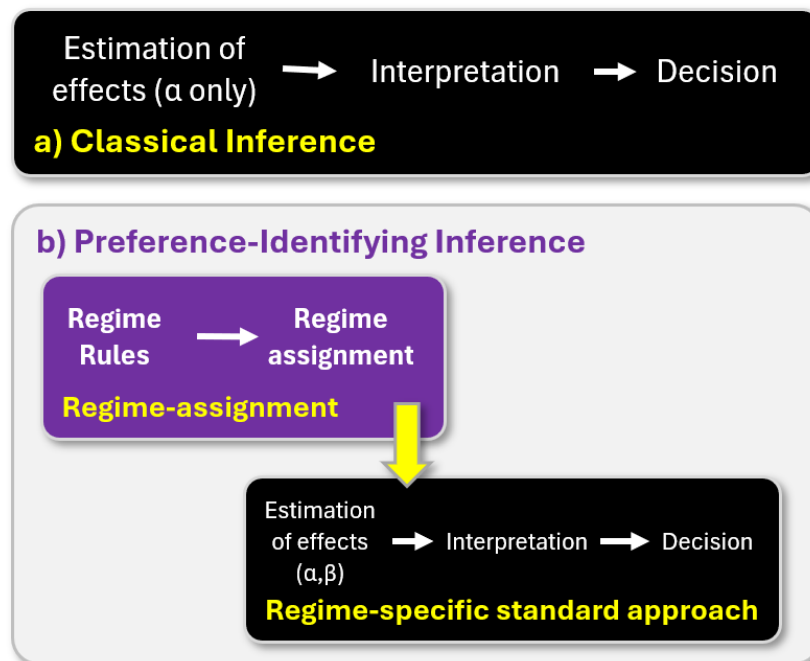
4. Discussion

When outcomes are jointly shaped by α and β , estimators alone do not suffice. Interpretation and design must follow the identified $(|\Delta\alpha|, \beta)$ structure. Trials must therefore resolve three questions: what α is, when β materially governs outcome, and when choice rather than biology becomes decisive.

4.1 Interpretation and Regime-Guided Trial Design

Interpretation follows regime assignment. Classical randomized inference moves from estimation to interpretation. This framework inserts a prior condition: regime assignment must first resolve the identified $(|\Delta\alpha|, \beta)$ structure (Figure 3a,b).^{24–26} Preference-identifying trials are therefore classified by prespecified rules, not estimates alone. Estimation serves the assigned structure rather than defining it. Under the Interpretation Rule (R1), interpretive authority turns on the joint magnitude of $|\Delta\alpha|$ and β , and design, power, and measurement must preserve the conditions for identifying that structure (R2–R4).^{16,23} Interpretation thus follows regime membership, not estimators alone.

Figure 3. From Classical Inference to a Regime-Based Framework for Causal Measurement of Patient Preferences



Panel (a) shows classical inference, in which interpretation follows directly from estimation. Panel (b) shows the regime-based framework, in which interpretation is conditioned on rule-based regime assignment; estimation informs identification but does not by itself determine interpretive meaning.

Regimes as inferential conclusions, not assumptions. Regime labels— α -dominant, mixed, β -dominant, and decision-limited—are inferential conclusions, not assumptions. No regime is fixed in advance; each is assigned only after $(|\Delta\alpha|, \beta)$ has been identified, or shown non-identifiable, on a common clinical scale.⁸ Design, however, must preserve the conditions for that assignment from the outset. If $|\Delta\alpha|$ may be small and β consequential, the Regime-Routing Rule (R2) requires a design that keeps preference effects identifiable. Under Dual Causation (A1) and the Regime Rule (A4), an α -only design can erase the very structure the trial must read.

The logic is analogous to randomization: effects are unknown in advance, yet design must still preserve identifiability.^{24,32} Under Dual Causation (A1), Identifiability Before Optimization (A5), and safeguards D1–D3 (see Table S1b in Supplement S1), that means preserving separation of biological treatment and preference-mediated choice effects. In Chalmers, divergence was visible but not decomposable, and the result was ambiguity.⁷ In SPORT, causal identification of α together with observed choice made decomposition into $|\Delta\alpha|$ and β possible.^{1,2,5,6} Regimes therefore guide design by risk rather than prediction: identifiability comes first, and only then can regime assignment proceed.

Implications for trial design. Tables 3 and 4 show that design is regime-routed rather than interchangeable. In α -dominant settings (B1), standard randomized designs suffice; in mixed settings (B2), randomization must be preserved, but inference must carry both α and β .^{25,26,54} In β -dominant settings (B4), adherence enforcement and crossover suppression are not neutral, because they condition the trial on an α -dominant estimand under standard IV logic^{8,27} and can obscure the preference-mediated structure to be identified. In decision-limited settings (B3), the task is stronger identification—through better instruments, improved elicitation, or revised power—before the structure can be read. In mixed and β -dominant regimes, suppressing uptake variation may increase α -centric rigor but block β -identification.^{5,6,8,27}

4.2 Empirical and Paradigmatic Considerations

Reinterpreting SPORT. SPORT (lumbar disc herniation) is best understood as an early β -dominant (B4) trial that α -centric interpretation could not fully read. High crossover and nonadherence were treated as threats to validity rather than as identifying variation, and the trial was therefore judged inconclusive.^{1–3} Yet SPORT satisfied the conditions for β -identification: randomized assignment identified α under nonadherence, while observed treatment received preserved the contrast needed to quantify β .^{5,6,8} What SPORT lacked were prespecified regime rules once $|\Delta\alpha|$ proved small; identifiable preference-mediated structure was therefore forced back into α -only logic. SPORT thus serves here as an empirical anchor of β -dominant structure, not yet as a fully β -identifying trial.

Other β -dominant clinical settings. The same regime structure recurs where biological contrast narrows while burden, trade-offs, and enacted choice continue to shape outcome. In nonvalvular atrial fibrillation, rhythm control confers no survival advantage, yet symptoms, harms, hospitalizations, and anticoagulation still carry consequence, placing the decision in mixed or β -dominant territory.^{36,49} In acute appendicitis, antibiotics are noninferior to appendectomy in the short term, but later surgery and complication risk still vary, so choice continues to shape outcome and the regime turns β -dominant.⁴⁷ The same structure plausibly extends to lumbar disc herniation, localized prostate cancer, stable coronary disease, benign prostatic hyperplasia,⁷ knee osteoarthritis, peripheral artery disease, symptomatic fibroids, obstructive sleep apnea, low-risk thyroid cancer, and multimorbidity deprescribing pathways, as detailed in Supplement S11.

In these settings, α -first interpretation no longer settles the decision, even when biological equipoise remains clinically coherent. What matters is not only whether this structure recurs, but whether trial design can preserve it as identifiable evidence.

Preference-aware designs do not identify preference effects (β). Typical patient-preference trials treat preference as a design constraint used to preserve α .^{25,26,54} Within such models, preference is handled as a threat to estimation, not as causal structure to be recovered. The design may accommodate preference in practice, yet the estimand remains α -centric: preference is not defined as an outcome-scale effect.

This resolves a common conflation: accommodating preference is not the same as identifying β . Pragmatic trials may amplify preference in practice yet still blur α and β in inference, leaving the relevant structure present but unidentified.^{4,40–42} Shared decision making, by contrast, is normative rather than causal. Without β measured on an outcome scale, it cannot distinguish choice-driven from biologically governed settings.^{31,61}

Trial design must keep α and β identifiable together (A5, Identifiability Before Optimization; R2, Regime-Routing).^{8,27} Preference trials, pragmatic trials, and shared decision making all engage preference, but none by itself separates biology from choice on the outcome scale.^{5,6} What is new is not acknowledgment of preference, but β as an estimand interpreted by regime. Once the trial enters a β -dominant regime, preference-governed inference is required.

Evolution of β across trial generations. To make the novelty explicit, β can be placed within the lineage of trial paradigms, not just within SPORT, as summarized by Table 5. SPORT marks the transitional case: randomization plus observed uptake made α and β separable in principle, but neither was prespecified nor interpretable under α -centric conventions.^{1,2,5,6} β -identifying trials, as outlined in this paper, complete that arc by prespecifying β as a causal

estimand and coordinating axioms (A), rules (R), and methods (M) before interpretation begins.^{5,6}

Table 5. Progression of β Across Trial Generations

	Chalmers (1981)	Protect	SPORT	β -Quantifying Secondary Analyses	β -Identifying Trials (This Paper)
Primary aim	Compare designs	Compare treatments	Compare treatment strategies	Decompose divergence	Identify and interpret causal structure
β defined as estimand	No	No	No	Yes	Yes (prespecified)
Role of crossover	Bias	Nuisance	Analytic problem	Source of β	Primary identifying variation
Regime classification possible	No	Implicit	Implicit	Retrospective	Prospective & rule-mediated
Epistemic status	Structure recognized	Structure acknowledged	Structure identifiable but unread	Structure decomposed <i>post hoc</i>	Structure anticipated, identified, interpreted

Historical progression in how clinical trials handle preference-mediated effects. Across generations, the role of preference evolves from unstructured divergence to an explicit causal estimand: crossover moves from bias (Chalmers)⁷ to nuisance (Protect)^{38,39} to analytic signal (SPORT)^{1,2,5,6} to primary identifying variation (β -identifying designs); and regime classification advances from absent to implicit to retrospective to fully prospective and rule-governed. Full versions in Supplement S8 (Tables S5a and S5b).

4.4 Direct Implications of β -Identification

Strengthen evidence-based medicine. Once β is identified, attenuation becomes structure rather than signal loss. Trials that are equivocal under α -centric analysis separate into α -dominant and β -influenced cases, so apparent ambiguity resolves into biology versus choice. By making explicit where α -centric inference applies, evidence-based medicine is strengthened.^{4,22,23}

Separate biological generalizability from contextual dependence. Identification of β separates invariant biological effects from context-dependent treatment uptake. α reflects biological contrast and is transportable across settings, whereas β varies with the structure of choice. External validity is no longer a single property of the trial; it separates into what generalizes (α) and what varies with context (β).^{62,63}

Policy Levers Follow β . Identification of β shifts the location of intervention. Where β is small, outcomes are governed by biological contrast and action lies in treatment selection within established evidence-based frameworks. Where β is substantial, outcomes depend on choice formation, and the operative levers shift to access, framing, literacy, and decision support.^{31,61,64} The object of intervention therefore moves with β rather than remaining fixed at the level of treatment alone.

Adherence Follows Treatment–Choice Coupling. β renders adherence as structure rather than noise. Differences in uptake reflect interaction between treatment and choice, not deviation from protocol. A treatment that induces higher uptake under patient choice yields a larger realized effect even at similar biological efficacy, while low uptake under an alternative defines a clinically consequential compliance gap. Adherence therefore defines realized effectiveness, not just protocol deviation; with β identified, it becomes measurable and therefore improvable.^{8,27,32}

Decompose sources of variation. Clinical variation arises from multiple sources, including supply, local practice patterns, and clinical judgment, as documented in work such as the Dartmouth Atlas.^{64–68} Without separation, these components are conflated and attributed to inefficiency or uncertainty. Identification of β isolates variation attributable to patient-mediated choice, distinguishing it from variation arising from provider behavior or system structure.^{5,6} Non-preference variation can then be identified and, where appropriate, reduced, rather than obscured within undifferentiated variation.^{4,22,23,40–42}

4.4 Broader Implications and Applications

Enabling N-of-1 inference. N-of-1 trials address the individual case when population averages no longer decide, but they still require two conditions: sufficient response stability and enough within-person contrast to register difference.^{32,59} In preference-sensitive care, the threshold shifts. Because choice is itself causal (A1), the signal that resolves the decision may lie less in α than in β .^{5,6}

In β -dominant regimes— $|\Delta\alpha|$ not decisive, β materially consequential (A3)— α -first ranking cannot resolve the choice (R1).²⁸ Regime-Routing (R2) therefore points toward β -identifying N-of-1 inference, whose task is to preserve separation of α and β on a common outcome scale.^{5,6,8} That does not remove the usual constraints of N-of-1 design—power, stability, and carryover remain—but when β is large it may become necessary, even if not sufficient. Without β identification, N-of-1 cannot support interpretable preference-sensitive decisions: preference-mediated choice collapses into undifferentiated treatment effect or noise, and the design becomes a label rather than an inference.

Real-world trials and β -identifying inference. Real-world trials, typically pragmatic or embedded in routine care, extend randomized research into ordinary clinical practice.^{55–58,69–}

⁷⁴ Yet real-world execution does not by itself identify β . Applicability can increase while uptake heterogeneity remains causally unread. Their value is therefore conditional: in mixed and β -dominant settings, uptake, switching, persistence, and patient-valued trade-offs may carry identifying information rather than mere variation. Under those conditions, β -identifying inference adds what real-world design alone does not: prespecified co-estimands, neutrality conditions for preference measurement, and reporting that preserves the causal structure generated in practice. Real-world trials supply the setting; β -identifying inference supplies the discipline.

Implications for AI decision support. The regime map in Figure 2 governs not only trial interpretation, but the legitimacy of clinical AI recommendation. When biological contrast is decisive (α -dominant, B1), α -based ranking remains legitimate because biology discriminates. When β is clinically consequential—as in mixed and β -dominant settings (B2, B4)—or when evidence is insufficient to decide (B3), ranking loses legitimacy and AI must defer pending explicit preference elicitation or stronger identification.¹³

This shift also determines how optimization is permitted. In α -dominant settings, optimization proceeds through improved prediction and treatment selection. Outside α -dominance, biological ranking is no longer admissible. Instead, once preferences are explicitly elicited, optimization must proceed through alignment and enactment— β -optimization—improving outcomes by increasing concordance between evidence, patient values, and what can be sustained in practice, rather than by selecting a superior treatment.¹⁴

Stakeholder ecosystem. β -identifying evidence will not arise from trial design alone. It requires an ecosystem in which sponsors, regulators, journals, and guideline bodies treat preference-mediated effects not as residual noise, but as part of the evidence itself.

The process begins with sponsors and trial designers. Protocols must prespecify estimands, power accordingly, and embed adaptation rules and identification diagnostics from the outset.^{4,29,32,75} Without such commitments, clinically important preference effects collapse into observational residue. Trials must therefore preserve uptake variation, enforce neutral elicitation, and maintain separation between biological contrast and enacted choice.^{5,6,8,27,55,56}

Journals and reporting standards form the second layer. If such trials are reported through α -centric conventions, their added structure becomes hard to see, compare, or reuse.

Reporting standards must therefore require co-reporting of the key causal contrasts, regime assignment, and identification diagnostics.^{4,24,55,75}

Regulators, guideline bodies, and health-technology assessment organizations determine how such evidence counts in approval, recommendation, value assessment, and reimbursement.^{21–23} Under near biological equipoise, guidance can no longer remain uniformly treatment-centric: the question shifts from biological superiority to which pathway best fits patient values and burden. Regime-conditional reasoning and patient-preference information may then extend existing frameworks. Otherwise, α -first frameworks will continue to measure biology while leaving the action-governing structure only partly visible.

4.5 Limitations and open problems

β -identifying inference remains conditional: it depends on valid identification, context-specific transport, stability of regime structure over time, and preference measurement that does not itself become intervention.

Identification limits. Separation of α and β depends on valid identification. Weak instruments and noisy observational contrasts do not merely reduce precision; they constrain estimation and regime classification.^{33,34} Limited compliance strains identifiability itself, placing pressure on Identifiability Before Optimization (A5) and on the power requirements in M3 and M6.^{8,27,32} Addressing those limits requires stronger instruments and more discriminating diagnostic thresholds.³²

Transportability of β . Preference effects do not transport invariantly, because β is carried by context-specific choices and shifts across settings and populations. Under Dual Causation (A1) and MCID discipline (A3), β must be remeasured rather than extrapolated. The consequence is practical as well as inferential: reporting must be stratified, and regime classification must remain context-specific.^{5, 62,63}

Dynamic regimes. As patients age, accrue comorbidity, or revise preferences, the joint structure ($|\Delta\alpha|$, β) can shift over time, and regime membership with it.^{9–11,76} Current formulations of the Interpretation Rule (R1) and Regime-Routing Rule (R2) are static and do not yet represent that movement. The next extension is therefore longitudinal: interpretation and design must follow ($|\Delta\alpha|$, β) as it evolves.^{30,77}

Measurement versus intervention. Preference elicitation must remain distinct from intervention. Once neutrality or timing is violated, β is no longer being measured alone; it is being altered, thereby challenging both Non-Interference of Elicitation (A6) and the

Measurement Rule (R4).^{15,21} As decision aids grow more sophisticated, valid measurement will therefore require explicit protocols and active auditing, not merely good intentions.^{61,78,79}

Together, these limits define a research agenda for β -identifying inference: stronger identification, better measurement, more careful transport, and designs able to track ($|\Delta\alpha|$, β) as it changes.

5. Conclusion

β -identifying trials complete randomized inference when biological contrasts narrow. By separating biological effects (α) from preference-mediated effects (β), the framework makes legible trials long judged equivocal, including SPORT, where crossover and nonadherence mark causal structure rather than trial failure.^{1,2} Once β enters view, the ($|\Delta\alpha|$, β) regime routes design, power, and clinical action, distinguishing settings in which standard designs suffice from those in which β -identifying designs become necessary.^{5,6} In β -dominant and mixed settings, care correspondingly shifts toward preference-concordant and, where feasible, N-of-1 approaches.

Classic randomized inference flourished where α dominated—acute disease, cleaner biology, fewer competing risks. As complexity rises through aging, multimorbidity, and polytherapy, $|\Delta\alpha|$ narrows, and β more often becomes clinically consequential, shifting care toward mixed and β -dominant regimes and making α -dominant findings from selective trial populations more circumscribed and less transportable to real-world settings.

This framework therefore does not weaken randomized evidence. On the contrary, it shows α -centric inference to be a special, albeit important, case within a broader structure aligned with the mechanisms generating outcomes more generally, and specifically in preference-sensitive care. Because preference-sensitive decisions govern a substantial share of care and resource use,^{64,80} rendering choice visible is not an ethical flourish but a methodological requirement: patient choice must be treated as causal structure, not absorbed into unmodeled noise.

Acknowledgements

The original SPORT study (ClinicalTrials.gov NCT00000410) was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (U01-AR45444-01A1) and the Office of Research on Women's Health, National Institutes of Health.

List of abbreviations

Abbreviation	Definition
AFT	Accelerated Failure Time model; survival analysis model expressed on a clinical outcome scale
AFFIRM	Atrial Fibrillation Follow-up Investigation of Rhythm Management trial
α (alpha)	Biological treatment effect; causal effect of treatment (assignment or causal effect under nonadherence), estimated via ITT or LATE/CACE
β (beta)	Preference-mediated effect; causal effect of patient choice, defined as $\beta = \Delta\text{OBS} - \Delta\alpha_{\text{causal}}$
$\Delta\alpha$	Net biological contrast between treatment options
$\Delta\beta$	Change in preference-mediated effect across contexts (e.g., different patient groups)
ΔOBS	Observed (as-treated) outcome difference
2SLS	Two-Stage Least Squares; estimation method for instrumental variable analysis
B1–B4	Regime classifications: B1 (α -dominant), B2 (mixed), B3 (decision-limited), B4 (β -dominant)
CACE	Complier Average Causal Effect; IV-based estimate of treatment effect among compliers
CI	Confidence Interval; measure of statistical uncertainty
CPAP	Continuous Positive Airway Pressure; therapy for obstructive sleep apnea
D1–D3	Design constraints: instrument validity, elicitation non-interference, uncertainty discipline
E-value	Sensitivity metric quantifying strength of unmeasured confounding needed to explain an effect
F-statistic	Diagnostic statistic assessing instrument strength in IV models
FEMME	Trial comparing uterine artery embolization vs myomectomy for fibroids
GLM	Generalized Linear Model; statistical modeling framework
ISCHEMIA	International Study of Comparative Health Effectiveness with Medical and Invasive Approaches
ITT	Intention-to-Treat; effect of treatment assignment

IV	Instrumental Variable; method for causal identification under nonadherence
LATE	Local Average Treatment Effect; causal effect among compliers
MCID	Minimal Clinically Important Difference; threshold for clinical relevance
M1–M6	Methods set: identification, β estimation, diagnostics, uncertainty, measurement, joint power
ML	Medical Literacy; patient understanding affecting treatment choice and outcomes
MTOPS	Medical Therapy of Prostatic Symptoms trial
π (pi)	Compliance parameter; measure of treatment uptake variation and IV strength
PRPP	Partially Randomized Patient Preference design
RCT	Randomized Controlled Trial
R1–R5	Rules: interpretation, regime routing, power, measurement, complexity drift
SDM	Shared Decision Making
SE	Standard Error; measure of estimator precision
SPORT	Spine Patient Outcomes Research Trial
 ProtecT	Prostate Testing for Cancer and Treatment trial

Declarations

Ethics approval and consent to participate

This study did not involve the collection or use of primary data from human participants or animals. It is a methodological manuscript based on Secondary analyses and conceptual framework development. Accordingly, ethics approval and consent to participate were not required.

Consent for publication

Not applicable. This manuscript does not contain any individual person's data, images, or identifying information.

Availability of data and materials

No new datasets were generated or analyzed during the current study. The work is based on previously published studies and methodological synthesis; all relevant sources are cited in the manuscript and Supplement.

Competing interests

The authors declare that they have no competing interests.

Funding

This work did not receive specific funding for the present methodology study. The original SPORT study referenced in this manuscript was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (U01-AR45444-01A1) and the Office of Research on Women's Health, National Institutes of Health.

Authors' contributions

O.G. conceived the study, developed the methodological framework, and drafted the manuscript. J.N.W. contributed to the development of the framework, interpretation of concepts, and critically revised the manuscript for important intellectual content. Both authors read and approved the final manuscript.

Use of Artificial Intelligence (AI) tools

The authors used AI-assisted tools solely for language refinement and editorial support in the preparation of the manuscript. No AI tools were used for hypothesis generation, study design, methodological development, data analysis, or interpretation. All intellectual content and conclusions are the responsibility of the authors.

References

1. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonoperative treatment for lumbar disc herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA*. 2006;296(18):2441–2450. doi:10.1001/jama.296.18.2441
2. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA*. 2006;296(20):2451–2459. doi:10.1001/jama.296.20.2451
3. McCormick PC. The Spine Patient Outcomes Research Trial results for lumbar disc herniation: a critical review. *J Neurosurg Spine*. 2007;6:513–520. doi:10.3171/spi.2007.6.6.1
4. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis*. 1967;20(8):637–48. doi:10.1016/0021-9681(67)90041-0
5. Gurel O, Rasmussen MF, Veginati V, Weinstein JN. AI Is Smart. Is It Wise? Quantifying the Effect of Patient Choice (β) on Physical Outcomes. *medRxiv [Preprint]*. 2026. Available from: <https://www.medrxiv.org/content/10.64898/2026.03.10.26348069v2> doi:10.64898/2026.03.10.26348069v2
6. Gurel O, Rasmussen MF, Veginati V, Weinstein JN. Treatment (α) and Patient Choice (β) Effects in Medicine: Quantifying the Causal Effect of Preference on Physical Outcomes. *Authorea [Preprint]*. 2026. Available from: <https://www.authorea.com/users/1037738/articles/1397573-treatment-%CE%B1-and-patient-choice-%CE%B2-effects-in-medicine-quantifying-the-causal-effect-of-preference-on-physical-outcomes> doi:10.22541/au.177446726.60210322/v3
7. Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D, et al. A comparison of results of randomized and non-randomized studies of surgical treatment for benign prostatic hyperplasia. *N Engl J Med*. 1981;305(23):1316–22. doi:10.1056/NEJM198111193052303
8. Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica*. 1994;62(2):467–75. doi:10.2307/2951620
9. Tinetti ME, McAvay GJ, Murphy TE, Gross CP, Lin H, Allore HG, et al. Mutual effects of multiple targeted medications on concurrent conditions in older adults with multiple diseases. *Ann Intern Med*. 2014;161(12):857–64. doi:10.7326/M14-1106
10. Pitkälä KH, Strandberg TE. Clinical trials in older people. *Age and Ageing*. 2022;51(5):afab282. doi:10.1093/ageing/afab282.
11. Fried TR, Tinetti ME, Agostini JV, Iannone L, Towle V. Health outcomes associated with polypharmacy in community-dwelling older adults. *J Am Geriatr Soc*. 2005;53(6):994–1000. doi:10.1111/j.1532-5415.2005.53315.x
12. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med*. 1987;317(3):141–5. doi:10.1056/NEJM198707163170304

13. Gurel O, Weinstein JN. From β -Blind to β -Aware AI for Preference-Sensitive Clinical Decisions: Achieving Non-Maleficence. Authorea [Preprint]. 2026. Available from: <https://www.authorea.com/users/1037738/articles/1398257-from-%CE%B2-blind-to-%CE%B2-aware-ai-for-preference-sensitive-clinical-decisions-achieving-non-maleficence> doi:10.22541/au.177499023.34125090/v2
14. Gurel O, Weinstein JN. From β -Aware to β -Optimizing AI for Preference-sensitive Clinical Decisions: Achieving Beneficence. Authorea [Preprint] 2026. Available from: <https://www.authorea.com/users/1037738/articles/1398258-from-%CE%B2-aware-to-%CE%B2-optimizing-ai-for-preference-sensitive-clinical-decisions-achieving-beneficence> doi:10.22541/au.177499025.52578495/v2
15. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: ethical principles and guidelines for the protection of human subjects of research. US Department of Health and Human Services, Office for Human Research Protections; 1979. Available from: <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>
16. Kass NE, Meeker-O'Connell A, Morain SR, Crane MA. Optimizing informed consent—A call to action. JAMA Health Forum. 2025;6(5):e250824. doi:10.1001/jamahealthforum.2025.0824
17. Moulton H, Moulton B, Lahey T, Elwyn G. Can consent to participate in clinical research involve shared decision making? AMA J Ethics. 2020;22(5):E365–71. doi:10.1001/amajethics.2020.365
18. Justia. Revised Code of Washington, RCW 41.05.033: shared decision making demonstration project—preference-sensitive care. Justia Law; 2021. Available from: <https://law.justia.com/codes/washington/2021/title-41/chapter-41-05/section-41-05-033>
19. Washington State Legislature. Revised Code of Washington, RCW 7.70.050: failure to secure informed consent—necessary elements of proof—emergency situations. Washington State Legislature; 2021. Available from: <https://app.leg.wa.gov/rcw/default.aspx?cite=7.70.050>
20. California State Legislature. California Health & Safety Code, § 24172: experimental subject's bill of rights. California Department of Justice, Office of the Attorney General; 1978. Available from: https://oag.ca.gov/sites/all/files/agweb/pdfs/research/safety_24172.pdf
21. Department of Health and Human Services. 45 CFR 46.116: general requirements for informed consent. Electronic Code of Federal Regulations (eCFR); [Accessed 12 May 2026] Available from: <https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.116>
22. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71–2. doi:10.1136/bmj.312.7023.71

23. Guyatt GH, Cairns J, Churchill D, Cook DJ, Haynes RB, Hirsh J, et al. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420–5. doi:10.1001/jama.1992.03490170092032
24. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. doi:10.1136/bmj.c332
25. Torgerson DJ, Sibbald B. Understanding controlled trials: what is a patient preference trial? *BMJ*. 1998;316(7128):360–1. doi:10.1136/bmj.316.7128.360.
26. Brewin CR, Bradley C. Patient preferences and randomised clinical trials. *BMJ*. 1989;299(6694):313–5. doi:10.1136/bmj.299.6694.313.
27. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58(1):21–9. doi:10.1111/1541-0420.00002
28. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407–15. doi:10.1016/0197-2456(89)90005-6
29. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC; 2020. Available from: <https://miguelhernan.org/whatifbook>
30. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560. doi:10.1097/00001648-200009000-00011
31. Charles C, Gafni A, Whelan T. Shared decision making in the medical encounter: what does it mean? *Soc Sci Med*. 1997;44(5):681–692. doi:10.1016/S0277-9536(96)00221-3
32. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. *Fundamentals of clinical trials*. 5th ed. Springer; 2015.
33. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica*. 1997;65(3):557–586. doi:10.2307/2171753
34. Stock JH, Yogo M. Testing for weak instruments in linear IV regression. In: Andrews DWK, Stock JH, editors. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press; 2005. p. 80–108. doi:10.1017/CBO9780511614491.006
35. Jin Z, Ying Z, Wei LJ. A simple resampling method by perturbing the minimand. *Biometrika*. 2001;88(2):381–90. doi:10.1093/biomet/88.2.381
36. Joglar JA, Chung MK, Armbruster AL, Benjamin EJ, Chyou JY, Cronin EM, et al. 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and management of atrial fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2024;149(1):e1–156. doi:10.1161/CIR.0000000000001193
37. Spertus JA, Jones PG, Maron DJ, et al; ISCHEMIA Research Group. Health-Status Outcomes with Invasive or Conservative Care in Coronary Disease. *N Engl J Med*. 2020;382(15):1408–1419. doi:10.1056/NEJMoa1916370

38. Hamdy FC, Donovan JL, Lane JA, Metcalfe C, Davis M, Turner EL, et al. 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *N Engl J Med*. 2016;375(15):1415-1424. doi:10.1056/NEJMoa1606220
39. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. *N Engl J Med*. 2016;375(15):1425-1437. doi:10.1056/NEJMoa1606221
40. Zuidgeest MGP, et al. Pragmatic trials and real world evidence: Paper 1. Introduction. *J Clin Epidemiol*. 2017;88:7–13. doi:10.1016/j.jclinepi.2016.12.023
41. Loudon K, et al. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350:h2147. doi:10.1136/bmj.h2147
42. Le-Rademacher JG, et al. From explanatory to pragmatic clinical trials. *Mayo Clin Proc*. 2023;98(8):1241–1253. doi:10.1016/j.mayocp.2023.04.012
43. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med*. 2008;358(8):794-810. doi: 10.1056/NEJMoa0707136
44. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis: primary results of the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *Spine (Phila Pa 1976)*. 2008;33(25):2744-52. doi: 10.1097/BRS.0b013e31818ee21c
45. Weinstein, J. N., Lurie, J. D., Tosteson, T. D., Hanscom, B., Tosteson, A. N. A., Blood, E. A., Birkmeyer, N. J. O., Hilibrand, A. S., Herkowitz, H., Cammisa, F. P., Albert, T. J., Emery, S. E., Lenke, L. G., Abdu, W. A., Longley, M., Errico, T. J., & Hu, S. S. (2007). Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *New England Journal of Medicine*, 356(22), 2257–2270. <https://doi.org/10.1056/nejmoa070302>
46. Lurie JD, Tosteson TD, Tosteson AN, Abdu WA, Zhao W, Morgan TS, et al. Surgical versus nonsurgical treatment for lumbar spinal stenosis of degenerative spondylolisthesis: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *Spine (Phila Pa 1976)*. 2008;33(25):2785-94. doi: 10.1097/BRS.0b013e3181893d5a
47. CODA Collaborative. A randomized trial comparing antibiotics with appendectomy for appendicitis. *N Engl J Med*. 2020;383(20):1907–19. doi:10.1056/NEJMoa2014320
48. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, et al. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med*. 2007;356(15):1503-16. doi: 10.1056/NEJMoa070829
49. Wyse DG, Waldo AL, DiMarco JP, et al; AFFIRM Investigators. A comparison of rate control and rhythm control in patients with atrial fibrillation. *N Engl J Med*. 2002;347(23):1825–1833. doi:10.1056/NEJMoa021328
50. Kirchhof P, Camm AJ, Goette A, Brandes A, Eckardt L, Elvan A, et al. Early rhythm-control therapy in patients with atrial fibrillation. *N Engl J Med*. 2020;383(14):1305-16. doi: 10.1056/NEJMoa2019422

51. Van Gelder IC, Groenveld HF, Crijns HJ, Tuininga YS, Tijssen JG, Alings M, et al. Lenient versus strict rate control in patients with atrial fibrillation. *N Engl J Med*. 2010;362(15):1363-73. doi: 10.1056/NEJMoa1001337
52. Hulley S, Grady D, Bush T, Furberg C, Herrington D, Riggs B, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women: Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *JAMA*. 1998;280(7):605-13. doi: 10.1001/jama.280.7.605
53. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321-33. doi: 10.1001/jama.288.3.321
54. Rücker G. A two-stage trial design for testing treatment, self-selection and treatment preference effects. *Stat Med*. 1989;8(4):477-85. doi:10.1002/sim.4780080411
55. Ford I, Norrie J. Pragmatic trials. *N Engl J Med*. 2016;375(5):454-63. doi: 10.1056/NEJMr1510059
56. Sox HC, Lewis RJ. Pragmatic trials: practical answers to “real-world” questions. *JAMA*. 2016;316(11):1205-6. doi: 10.1001/jama.2016.12151
57. U.S. Food and Drug Administration. Integrating randomized controlled trials for drug and biological products into routine clinical practice [Internet]. Silver Spring (MD): FDA; 2024 Feb [cited 2026 Jun 9]. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/integrating-randomized-controlled-trials-drug-and-biological-products-routine-clinical-practice>
58. NIH Pragmatic Trials Collaboratory. Rethinking clinical trials: a living textbook of pragmatic clinical trials [Internet]. Bethesda (MD): NIH Pragmatic Trials Collaboratory; 2026 [updated 2026 Jun 3; cited 2026 Jun 9]. Available from: <https://rethinkingclinicaltrials.org/>
59. Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid CH, Tate R, et al; CENT Group. CONSORT extension for reporting N-of-1 trials (CENT) 2015 statement. *BMJ*. 2015;350:h1738. doi:10.1136/bmj.h1738
60. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*. 2007;28(2):182-191. doi:10.1016/j.cct.2006.05.007
61. Elwyn G, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361-1367. doi:10.1007/s11606-012-2077-6
62. Pearl J, Bareinboim E. External validity: From do-calculus to transportability across populations. *Statist Sci*. 2014;29(4):579-595. doi:10.1214/14-STS486
63. Pearl J, Bareinboim E. Transportability of causal and statistical relations: A formal approach. In: *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI-11)*. 2011. p. 247-254. doi:10.1609/aaai.v25i1.7861
64. Wennberg JE. *Tracking medicine: a researcher's quest to understand health care*. Oxford: Oxford University Press; 2010.

65. Wennberg JE, Gittelsohn A. Small area variations in health care delivery. *Science*. 1973;182(4117):1102-8. doi: 10.1126/science.182.4117.1102
66. Wennberg JE. Practice variations and health care reform: connecting the dots. *Health Aff (Millwood)*. 2002;21(3):11-2. doi: 10.1377/hlthaff.21.3.11
67. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff (Millwood)*. 2002;21(Suppl):W96-W114. doi: 10.1377/hlthaff.21.suppl_1.w96
68. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Ann Intern Med*. 2003;138(4):273-87. doi: 10.7326/0003-4819-138-4-200302180-00006
69. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-world evidence—what is it and what can it tell us? *N Engl J Med*. 2016;375(23):2293-7. doi: 10.1056/NEJMs1609210
70. U.S. Food and Drug Administration. Integrating randomized controlled trials for drug and biological products into routine clinical practice [Internet]. Silver Spring (MD): FDA; 2024 Feb [cited 2026 Jun 9]. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/integrating-randomized-controlled-trials-drug-and-biological-products-routine-clinical-practice>
71. NIH Pragmatic Trials Collaboratory. Rethinking clinical trials: a living textbook of pragmatic clinical trials [Internet]. Bethesda (MD): NIH Pragmatic Trials Collaboratory; 2026 [updated 2026 Jun 3; cited 2026 Jun 9]. Available from: <https://rethinkingclinicaltrials.org/>
72. Mc Cord KA, Hemkens LG, Agarwal A, Perel P, Agarwal A, Ioannidis JP, et al. Clinical trials embedded in healthcare systems: improving efficiency and generalizability. *Clin Trials*. 2017;14(5):509-16. doi: 10.1177/1740774517719602
73. Fröbert O, Lagerqvist B, Erdbrügger W, Östlund O, James SK, Eriksson P, et al. Thrombus aspiration during ST-segment elevation myocardial infarction. *N Engl J Med*. 2013;369(17):1587-97. doi: 10.1056/NEJMo1308789
74. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci*. 2011;13(2):217-24. doi: 10.31887/DCNS.2011.13.2/npatsopoulos
75. International Council for Harmonisation (ICH). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. EMA/CHMP/ICH/436221/2017. Geneva: ICH; 2020.
76. Yang C, Vart P. Emulating target trials in older adults with multimorbidity. *Lancet Healthy Longev*. 2025;6(8):100750. doi:10.1016/j.lanhl.2025.100750
77. Kidwell KM, Almirall D. Sequential, multiple assignment, randomized trial designs. *JAMA*. 2023;329(4):336–337. doi:10.1001/jama.2022.24324
78. Stiggelbout AM, et al. Shared decision making: really putting patients at the center. *Patient Educ Couns*. 2012;90(2):144–150. doi:10.1016/j.pec.2012.05.002
79. Kunneman M. Shared decision making to design care that fits for each patient. *Nat Rev Dis Primers*. 2024;10:21. doi:10.1038/s41572-024-00522-0

80. Centers for Medicare & Medicaid Services. National Health Expenditure Accounts (NHEA). Baltimore: CMS; 2024 [cited 2026 Apr 26]. Available from: <https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [BetaIdentificationMethodsPaperSupplementv123Trials.docx](#)