The Methods of CER: Newer forms of randomized trials

Harold C. Sox, M.D., MACP

Director, Peer Review

The Patient-Centered Outcomes Research Institute

Professor of Medicine (emeritus)

Geisel School of Medicine at Dartmouth

Randomized trials

- Pragmatic trials
- Cluster trials
- Adaptive trials
- Non-inferiority trials

Pragmatic Trials

Efficacy Trials

- Also called explanatory trials
- Placebo-controlled; blinded
- Many exclusion criteria → narrow population
- Outcomes typically objective (BP, mortality)
- Tells you if the active intervention has an effect
- Designed to get the answer quickly and cheaply
- Applicability often poor

Pragmatic trials

- **Active control intervention (often "usual care")**
- **Ideally, take place in typical care settings**
- **Few exclusion criteria → typical population**
- **Outcomes typically objective events but may be symptoms**
- **Tells you if the active intervention has an effect in the real world; applicability is good.**

Sox HC, Lewis RJ. Pragmatic Trials: Practical Answers to "Real World" Questions. JAMA. 2016;315: 1205-1206

Explanatory vs. Pragmatic Trials

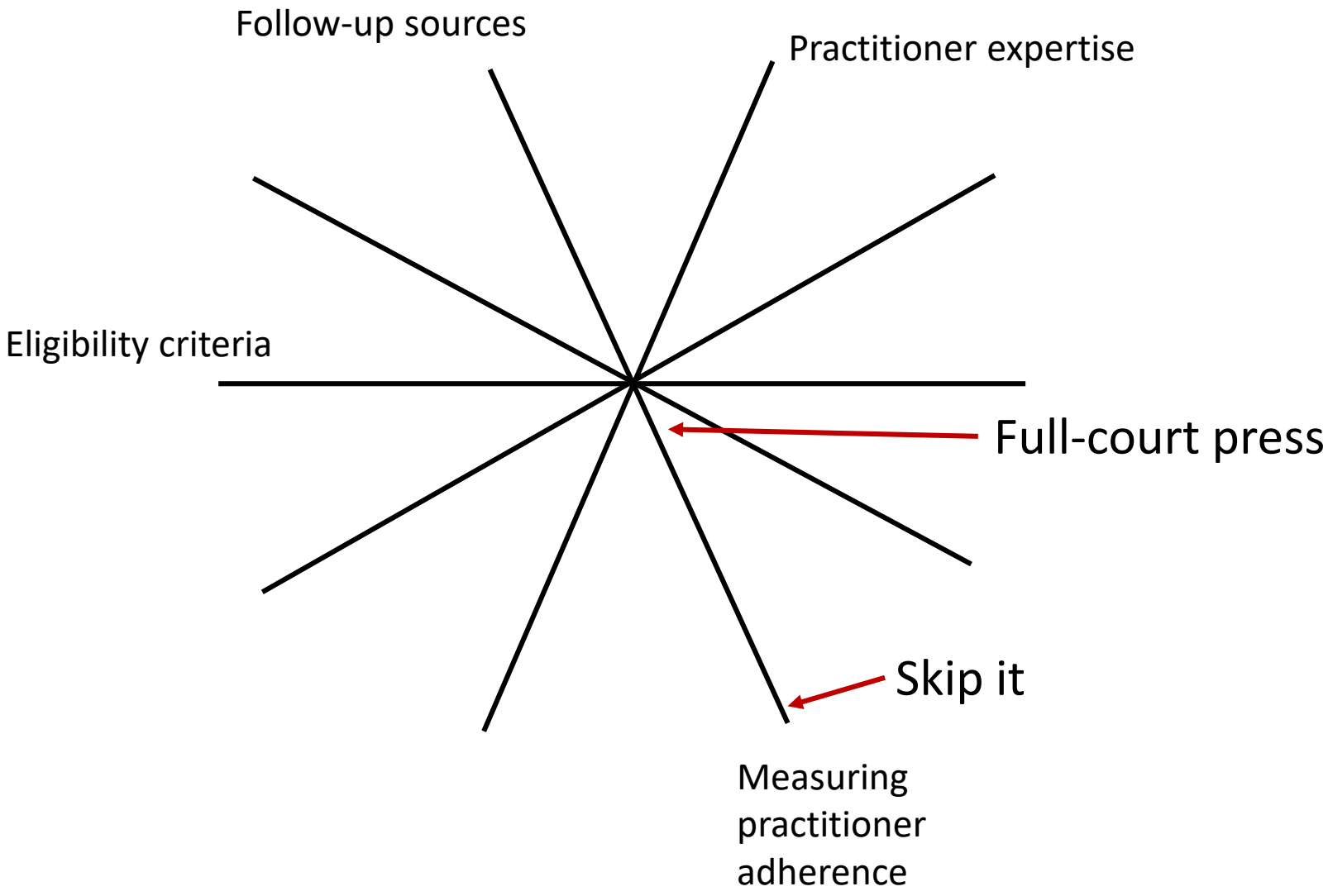
Factor	Pragmatic trial	Explanatory trial
Eligibility criteria	Everyone	Hi risk, likely to respond, hi adherence
Study protocol flexibility	Flexible; lots of leeway	Inflexible; detailed instructions;
Practitioner expertise	Full-range of practitioners and settings	Expert practitioners and settings
Comparator	“usual practice” or “best alternative”; lots of leeway	Not flexible; often placebo
Follow-up sources	Administrative data bases; EMR; death index	Many follow-up visits and data collection
Primary study outcome	Objectively measured, clinically meaningful; no special tests needed	Often requires special testing; surrogate for other outcomes

Thorpe KE et al. CMAJ; 2009;180;e47-e57

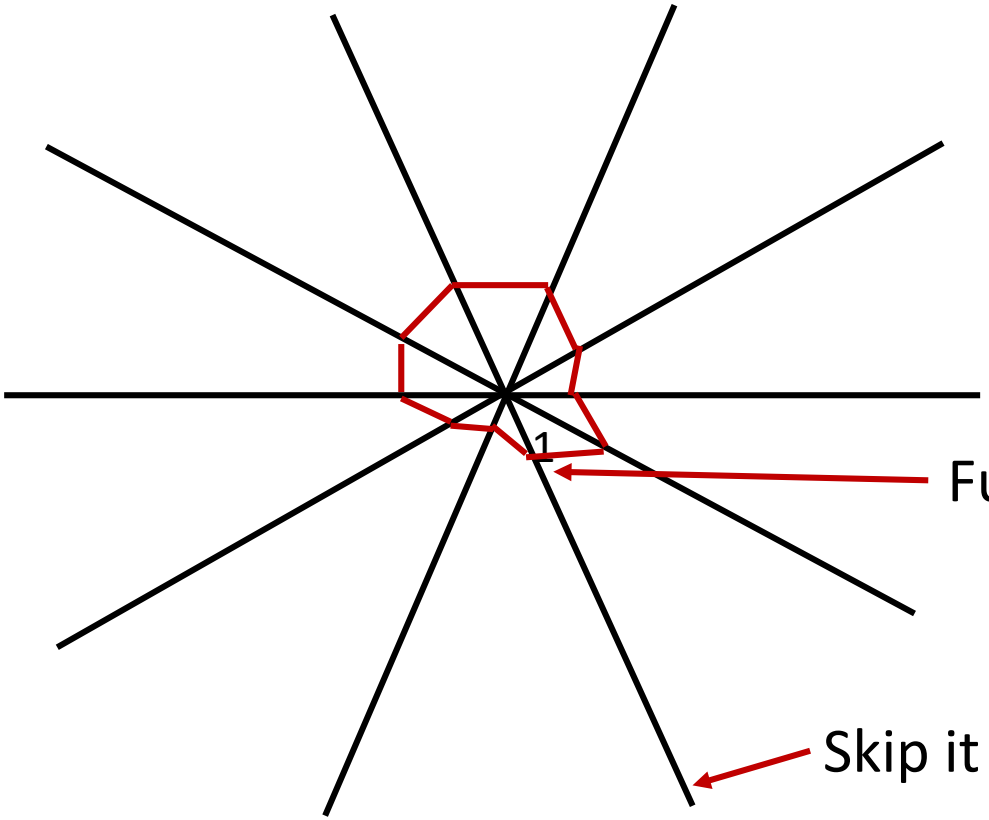
Explanatory vs. Pragmatic Trials

Factor	Pragmatic trial	Explanatory trial
Measuring patient adherence to intervention	Unobtrusive or no measurement; no effort to improve it	Monitored closely; use strategies to improve it
Measuring practitioner adherence to protocol	Unobtrusive or no measurement; no effort to improve it	Closely monitored
Analysis of primary outcome	All patients included (intention to treat)	Intent to treat but also per-protocol

Thorpe KE et al. CMAJ; 2009;180;e47-e57


Follow-up sources
Practitioner expertise
Eligibility criteria
Full-court press
Skip it
Measuring practitioner adherence

Explanatory Trial

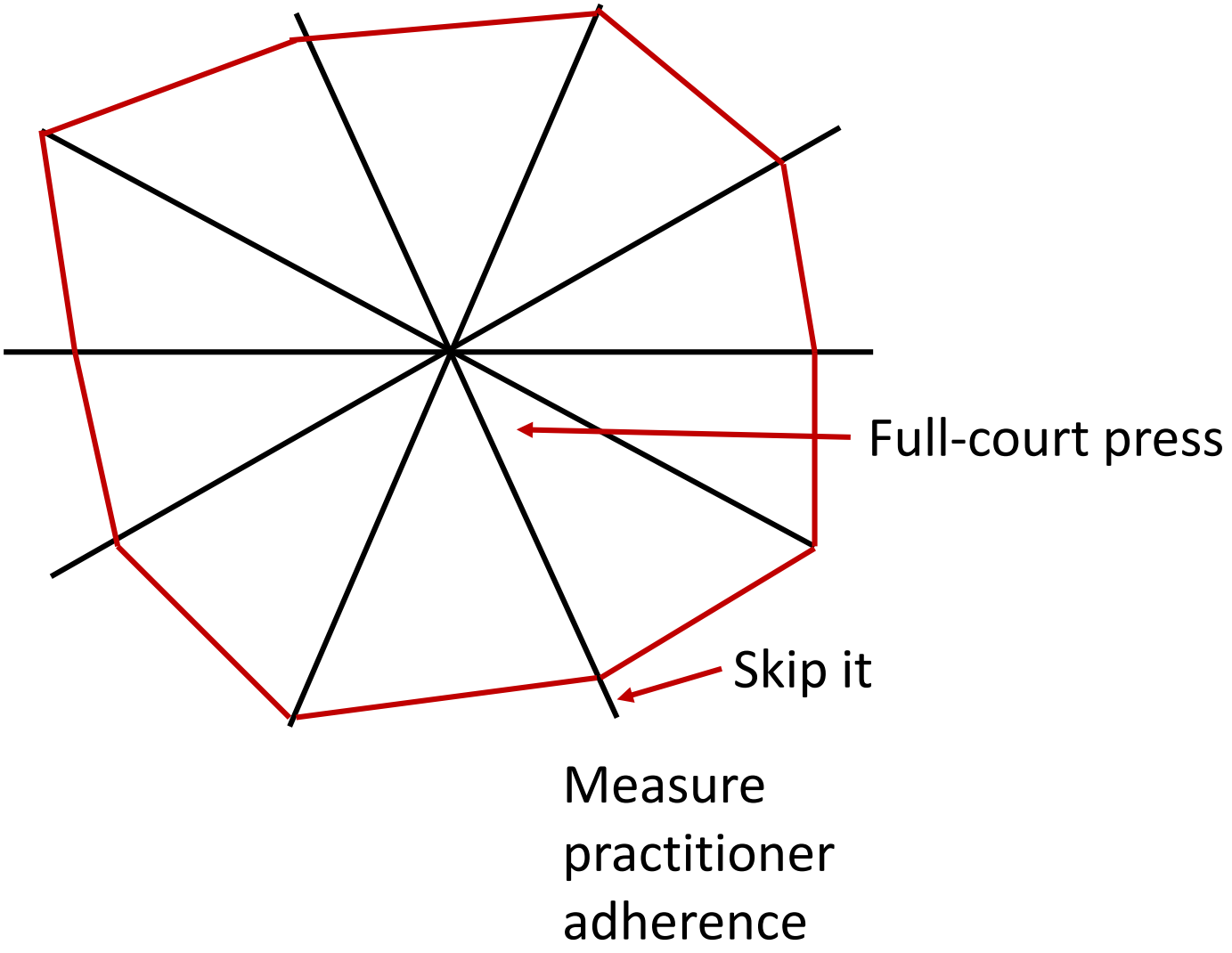


Full-court press

Skip it

Measure practitioner adherence

Pragmatic Trial



Mistakes in designing pragmatic trials

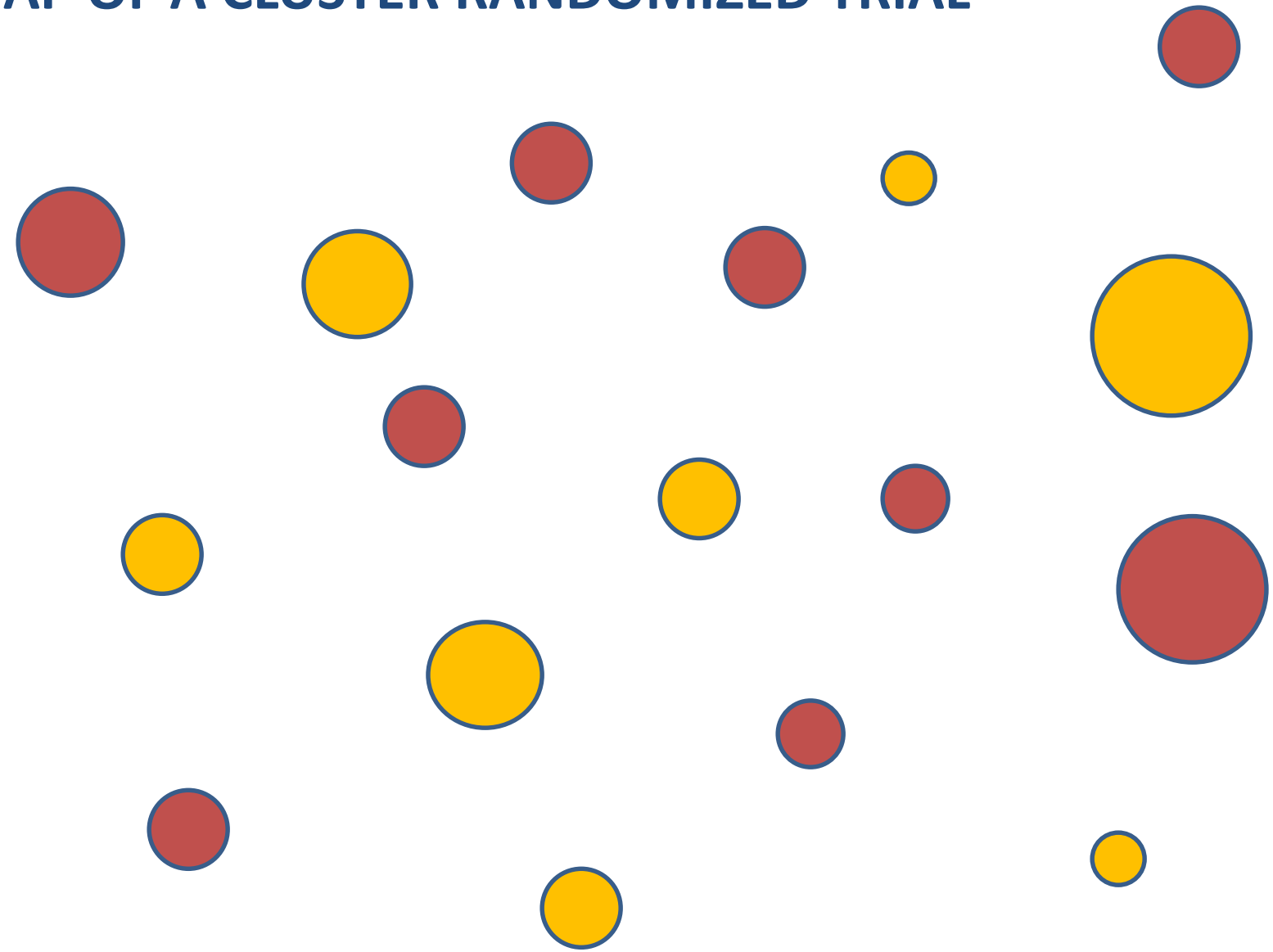
- With no measure of intervention uptake, hard to interpret a null result.
- Ceiling effects → hard to prove an effect.
- Relying on EHR/claims data for outcome → miss patient-reported outcomes.
- Hands-off approach → ↑missing data

Cluster Trials

Cluster trials

- **Instead of randomizing patients, randomize clinical sites.**

A MAP OF A CLUSTER RANDOMIZED TRIAL



Cluster trials

- **Advantage of randomizing by clinical site**
  - The best/only choice when the intervention alters patient care workflow.
  - Avoids contamination at the patient and provider level; everyone at a site gets the same intervention.
- **Disadvantages:**
  - Need lots of sites → complicated to run the trial
    - The ability of a study to detect a difference (power) depends on the number of sites more than cluster size.
  - Patients at a site tend to reflect practices of the site → less between-patient variation than if patients were independent of each other → confidence intervals too narrow → false-positive study result

Adjusting for clustering of outcomes by site

- **Must adjust for within-practice clustering**
  - Need larger sample size → adequately powered study
  - Measuring study outcomes → valid comparisons
- **Intra-class correlation coefficient is a key measure**

Stepped-wedge cluster randomized trial

- **Unidirectional randomized transition from control to active intervention.**
  - Randomly pick a site to implement the intervention; remaining sites are controls.
  - After X months, randomly pick a second site to implement the intervention; remaining sites are controls.
  - And so on until all sites have implemented the intervention

A stepped wedge cluster design

Site	January	Feb-ruary	March	April	May	June
6						
4						
3						
2						
1						

Intervention cluster

Control cluster

1. Everyone gets the intervention eventually
2. Within- and between-cluster comparisons strengthen statistical power
3. Can measure effects over time

Adaptive Trials

Why adaptive trials?

- They are suited to a fast-moving environment
  - Get early signals about ultimate outcomes
  - new drugs enter the marketplace during the trial. (e.g., Hepatitis C Rx)
  - Final results of a long trial may be irrelevant to current practice.
- Reliance on frequentist statistics (p values) can lead to a yes-no decision that
  - Oversimplifies
  - Ignores the pre-trial odds that one treatment would be superior

Bayes' Theorem

- **Basic principle: the interpretation of new information depends on what you thought beforehand.**
- **Derived from first principles; it is true.**
- **Used in test interpretation:**
  - Post-test odds = pre-test odds x likelihood ratio.
  - LR+= sensitivity/(1-specificity)
- **Used in adaptive trials**
  - Post-test odds = pre-test odds x Bayes factor
  - Bayes Factor: "D" is now = "treatment A is better" and "no D" is now "treatment B is better."

Adaptive Trials

- **Start with randomizing A vs. B.**
- **As results accumulate, use Bayes' theorem to calculate odds that A is better than B.**
- **If the odds that A>B reaches a pre-determined threshold, randomize fewer to B and start randomizing some to C.**
  - C is a new treatment with high odds of being better than A.
- **Treatment response heterogeneity**
  - If patients in a subgroup do much better with C, preferentially randomize them to A (vs. C).
- **The threshold probability for switching can reflect patients' feelings about the outcome states they might experience.**

Advantages of adaptive trials

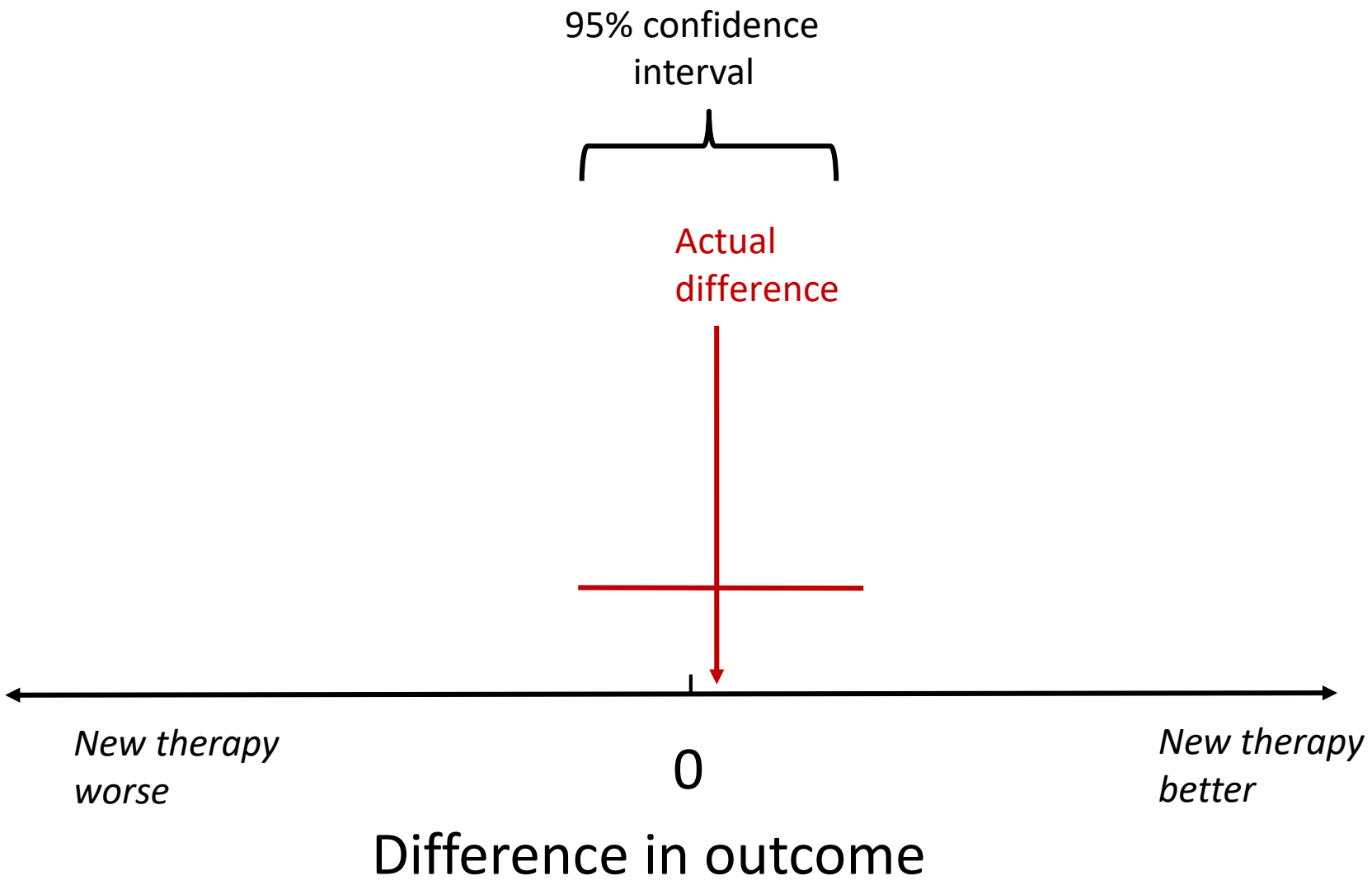
- **Efficient: gets to an answer sooner**
  - Fewer patients randomized to inferior treatment
  - Focuses the experimental resources (money, patients) on the best interventions
- **Maximizes patients' chance of being randomized to treatment they should prefer.**
- **Ideal in a fast-moving marketplace**

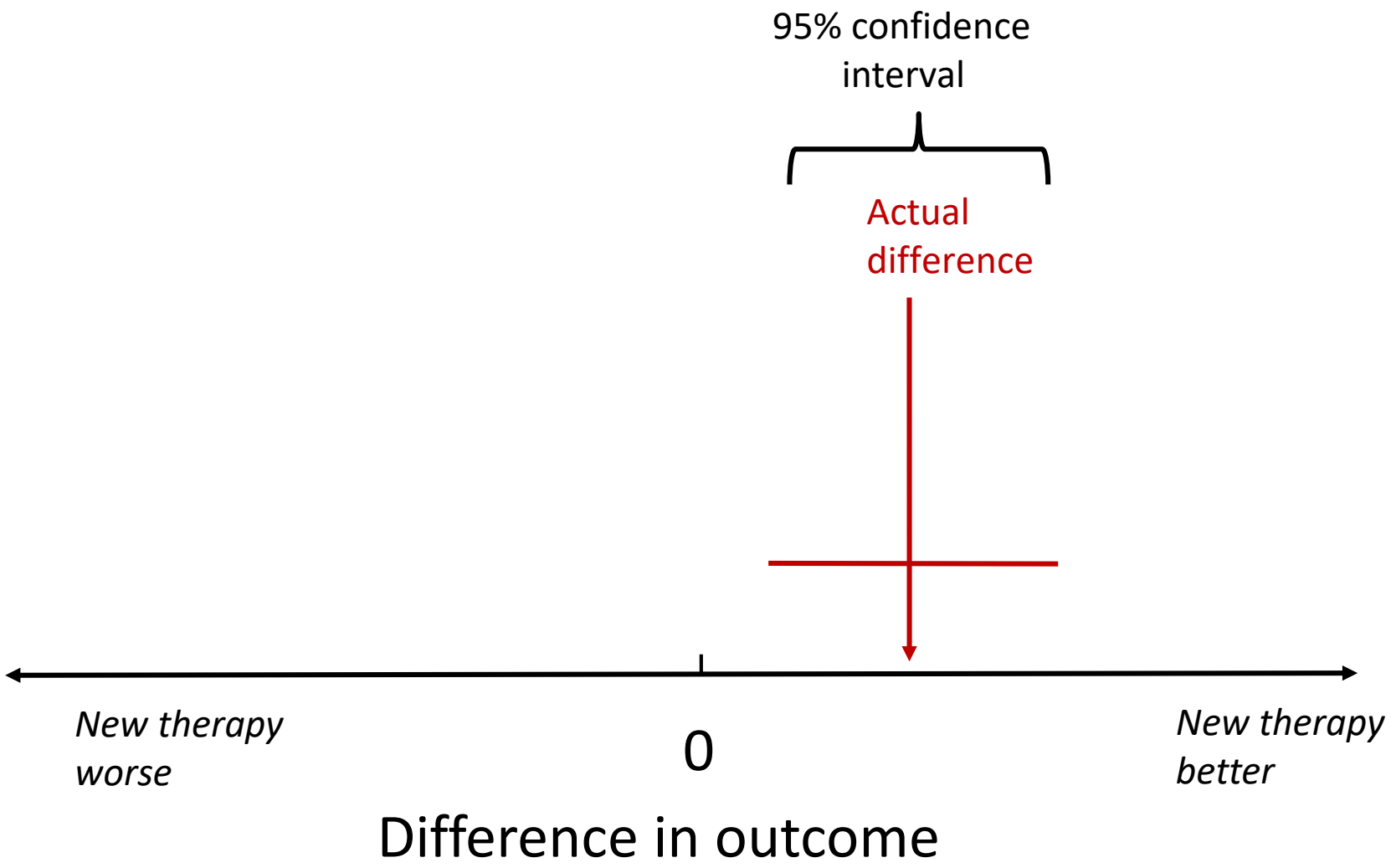
Non-inferiority trials

Why do a non-inferiority study?

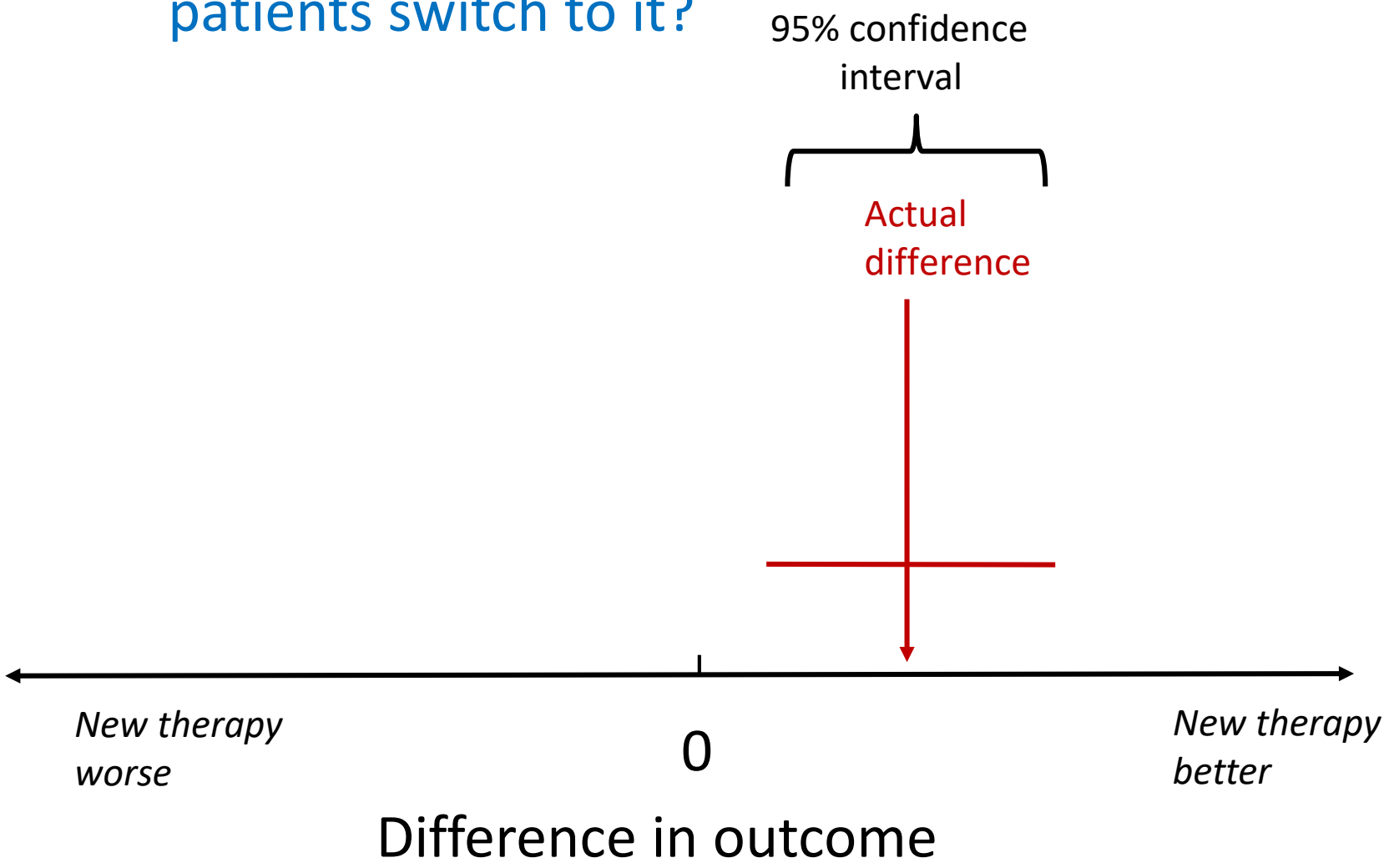
- **Is a cheaper treatment no worse than an accepted treatment?**
- **A head-to-head comparison could answer this question**
- **The cheaper treatment could be superior to, inferior to, or non-inferior to the established treatment.**
  - Cheaper and equivalent is good!

The Playing Field




95% confidence interval
Actual difference
New therapy worse
0
New therapy better
Difference in outcome

It's better but should patients switch to it?



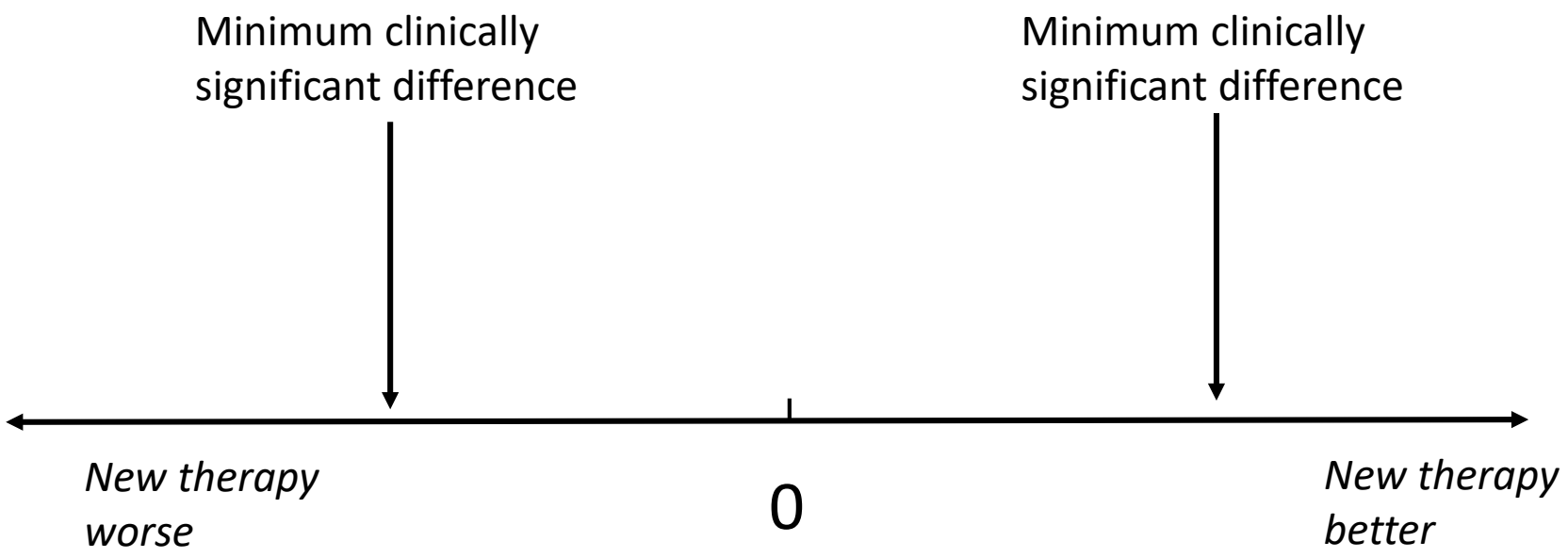
- What about a “positive” study comparing A to B? The effect is real, but the comparison can have several meanings.
  - Better: switch to the new therapy
  - Not worse (non-inferior): your call
  - Probably not inferior but could be better: your call
  - Probably not inferior but could be worse: your call
  - Worse: stay with the old therapy

The minimal clinically significant difference

- **The MCID is a key concept in the interpretation of trial results.**
  - The MCID is the smallest difference between the study outcome in the two study arms that is large enough to warrant a change to the treatment with the better outcome.
- **Determining the MCID**
  - In many trials, the MCID is an arbitrarily chosen fraction of the standardized difference.
  - An area for research

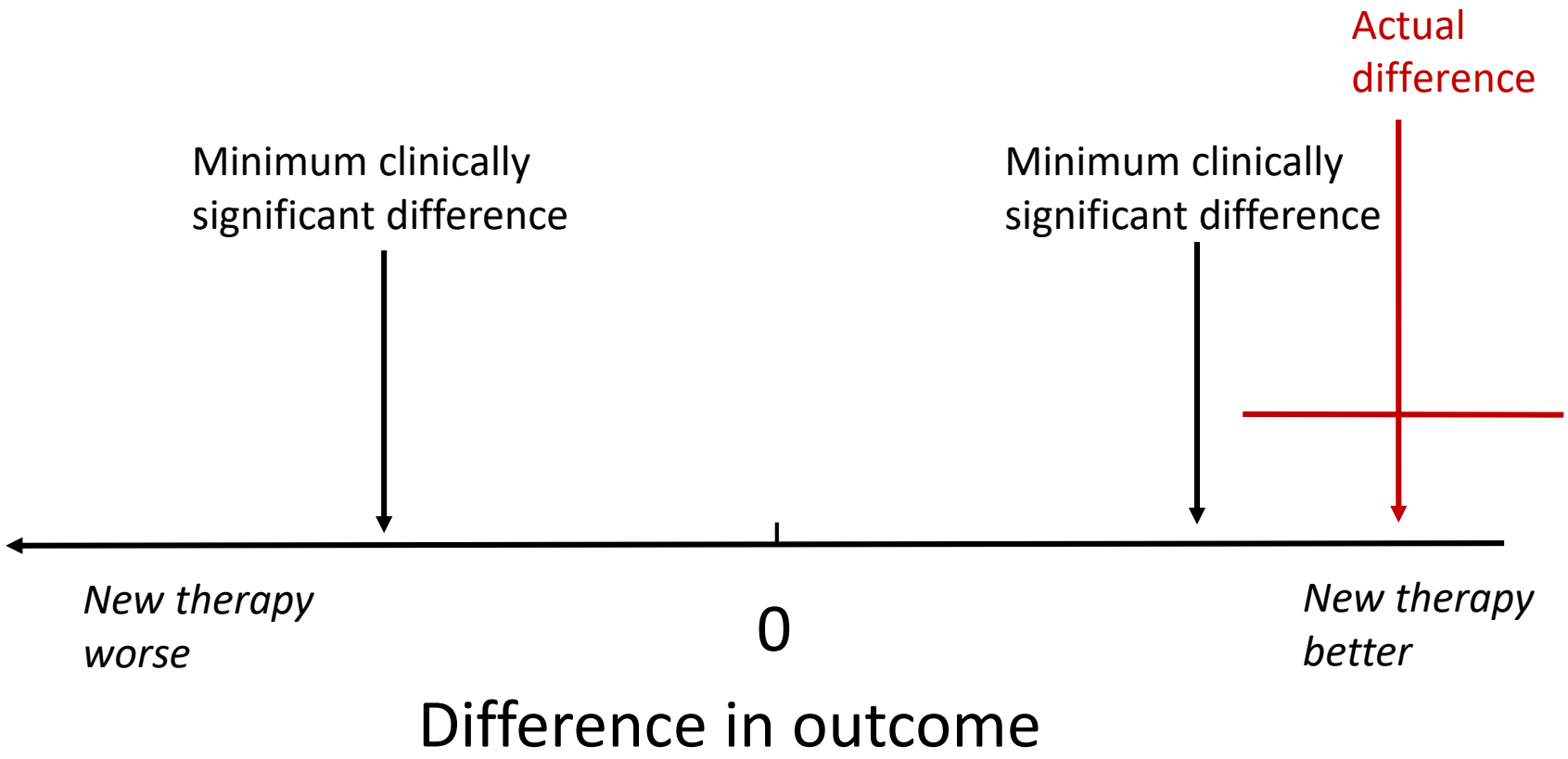
A Better Playing Field

The "minimum clinically significant difference:"
A difference just large enough to justify switching to the better intervention

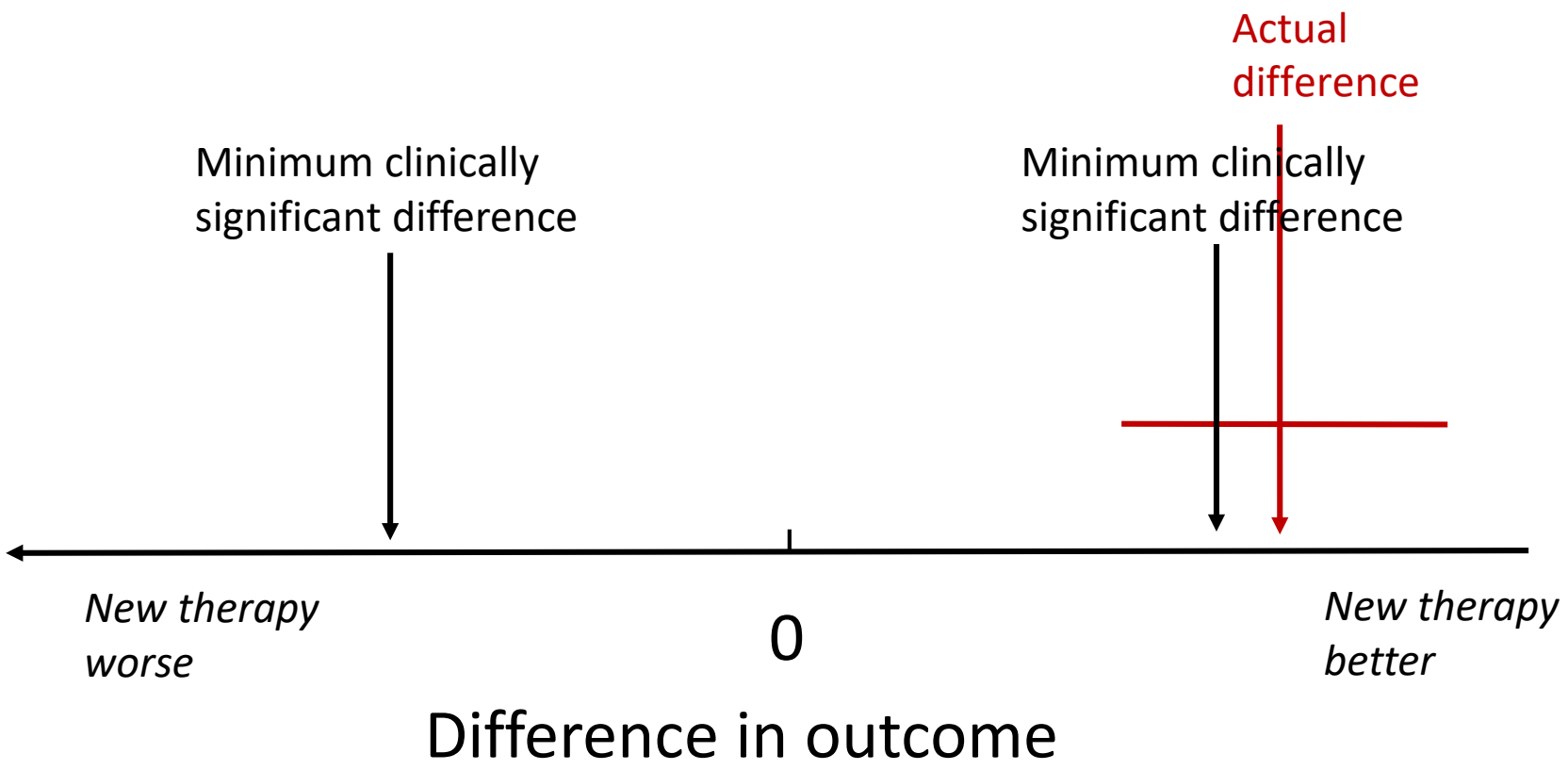


Difference in outcome between new and old therapy

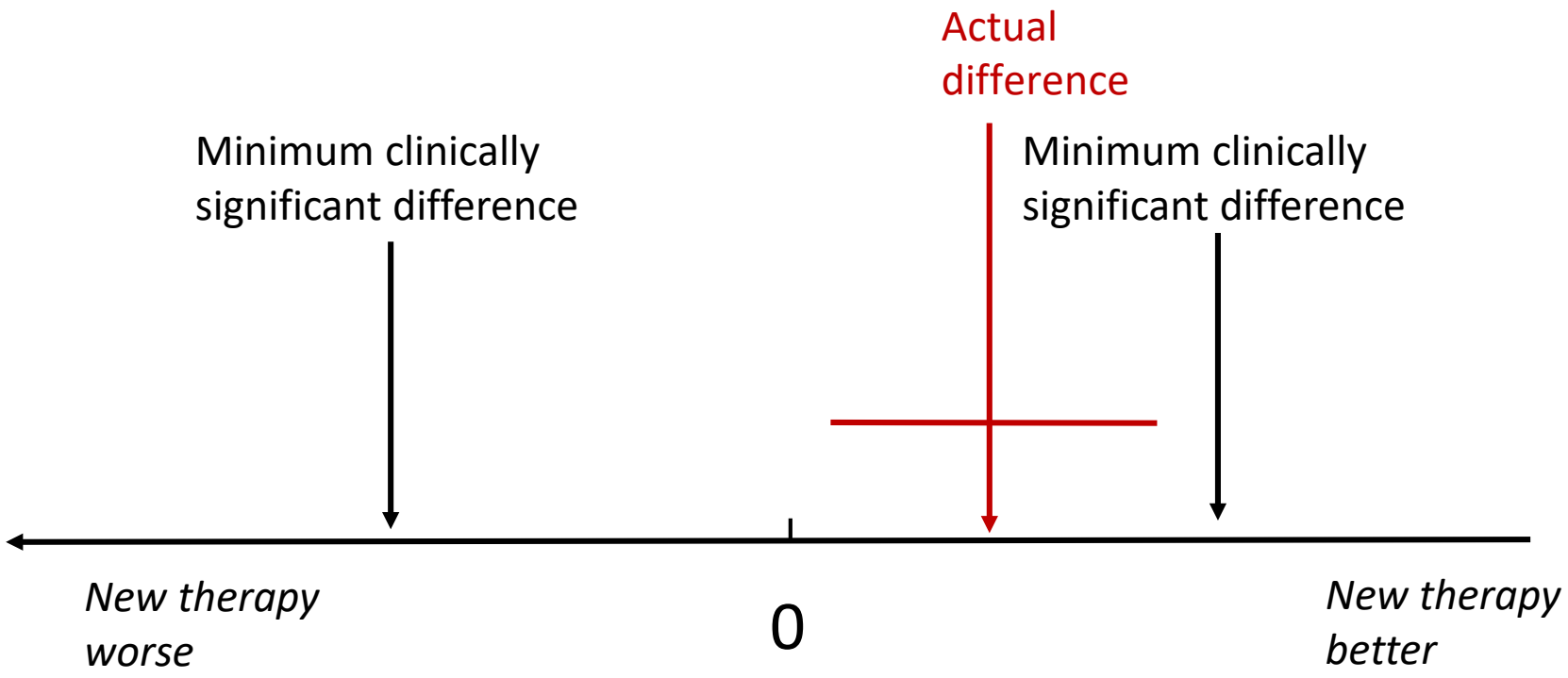
Statistically significant; New therapy better.



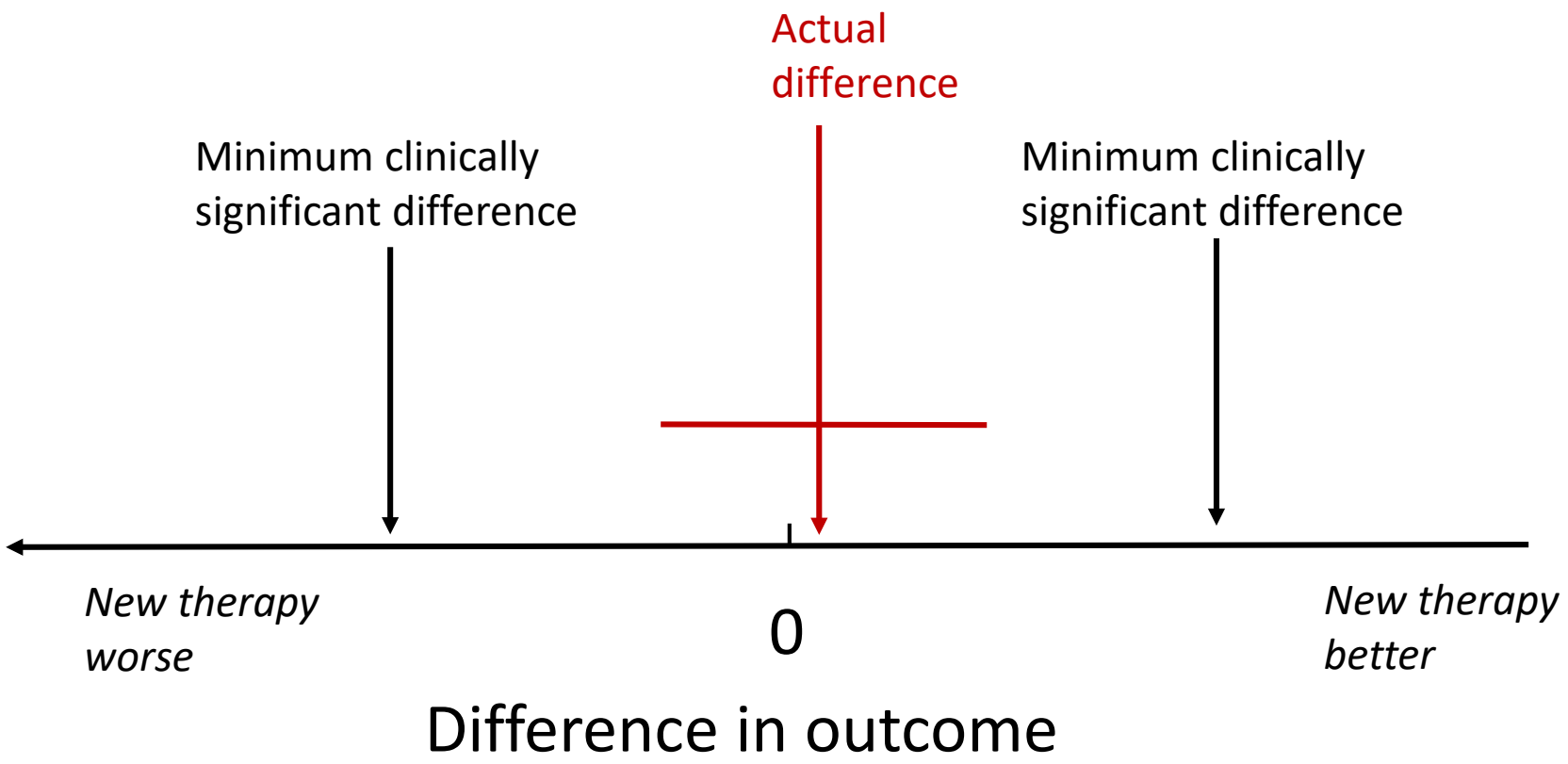
Statistically significant; Uninterpretable result for non-inferiority



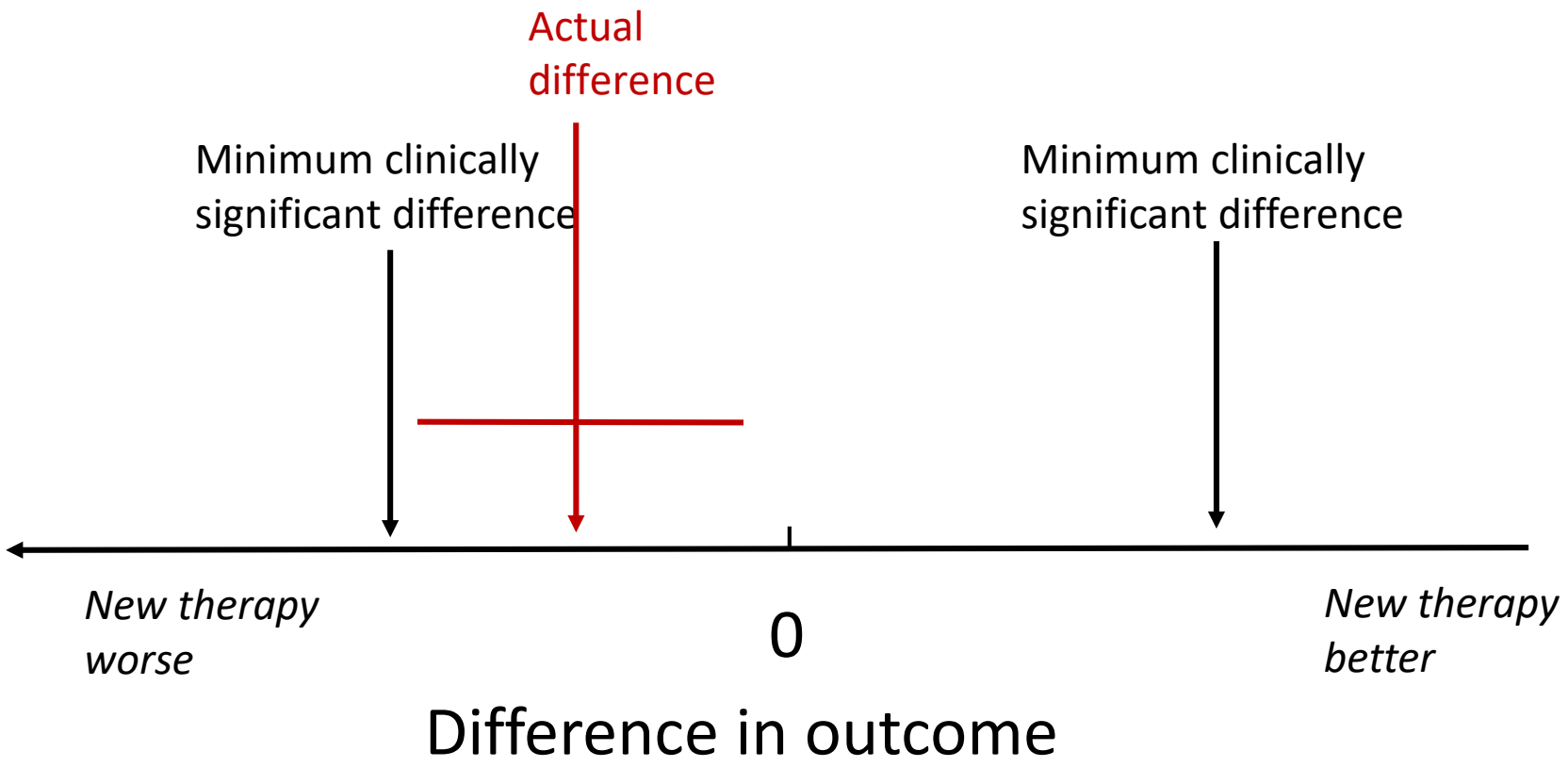
Statistically significant;
Not better or worse; a tie;
The new therapy is non-inferior



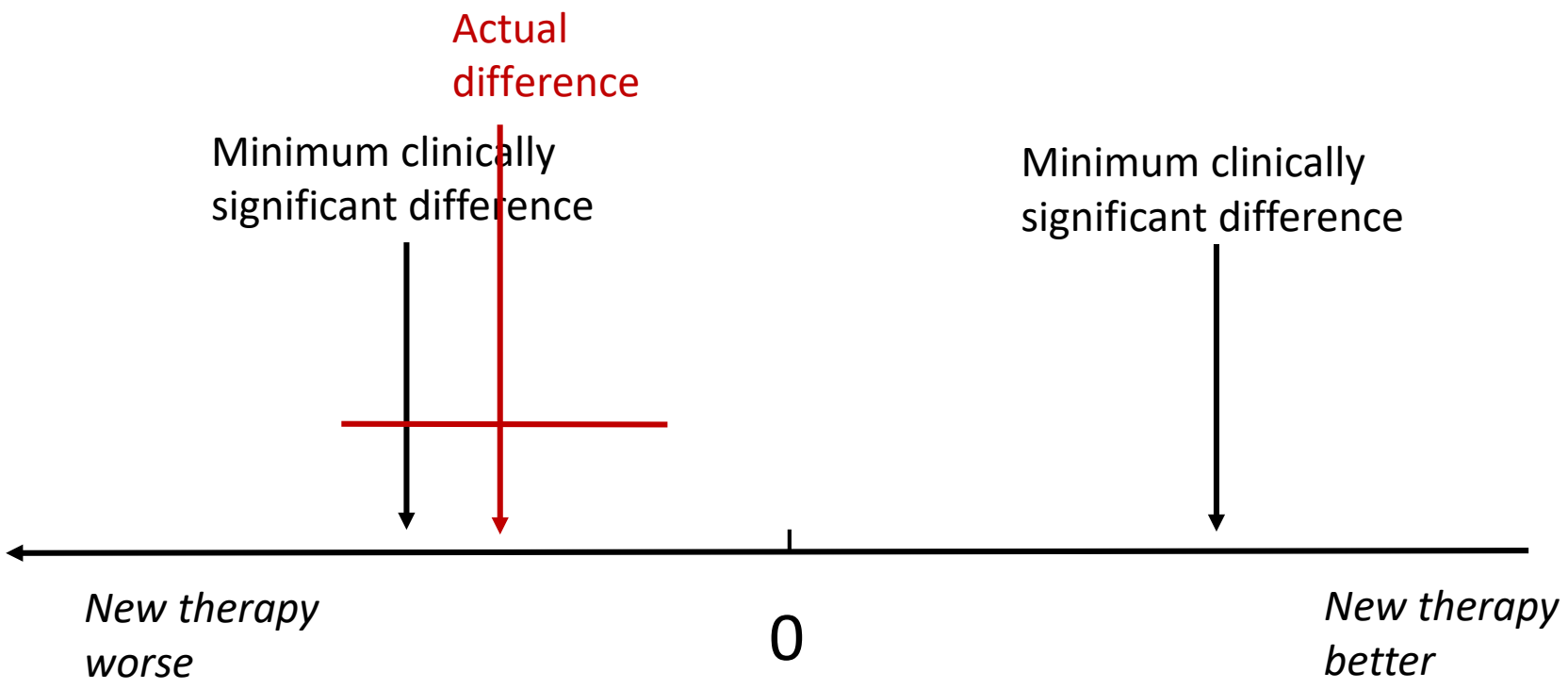
Statistically not significant; Not better or worse; a tie; Non-inferior



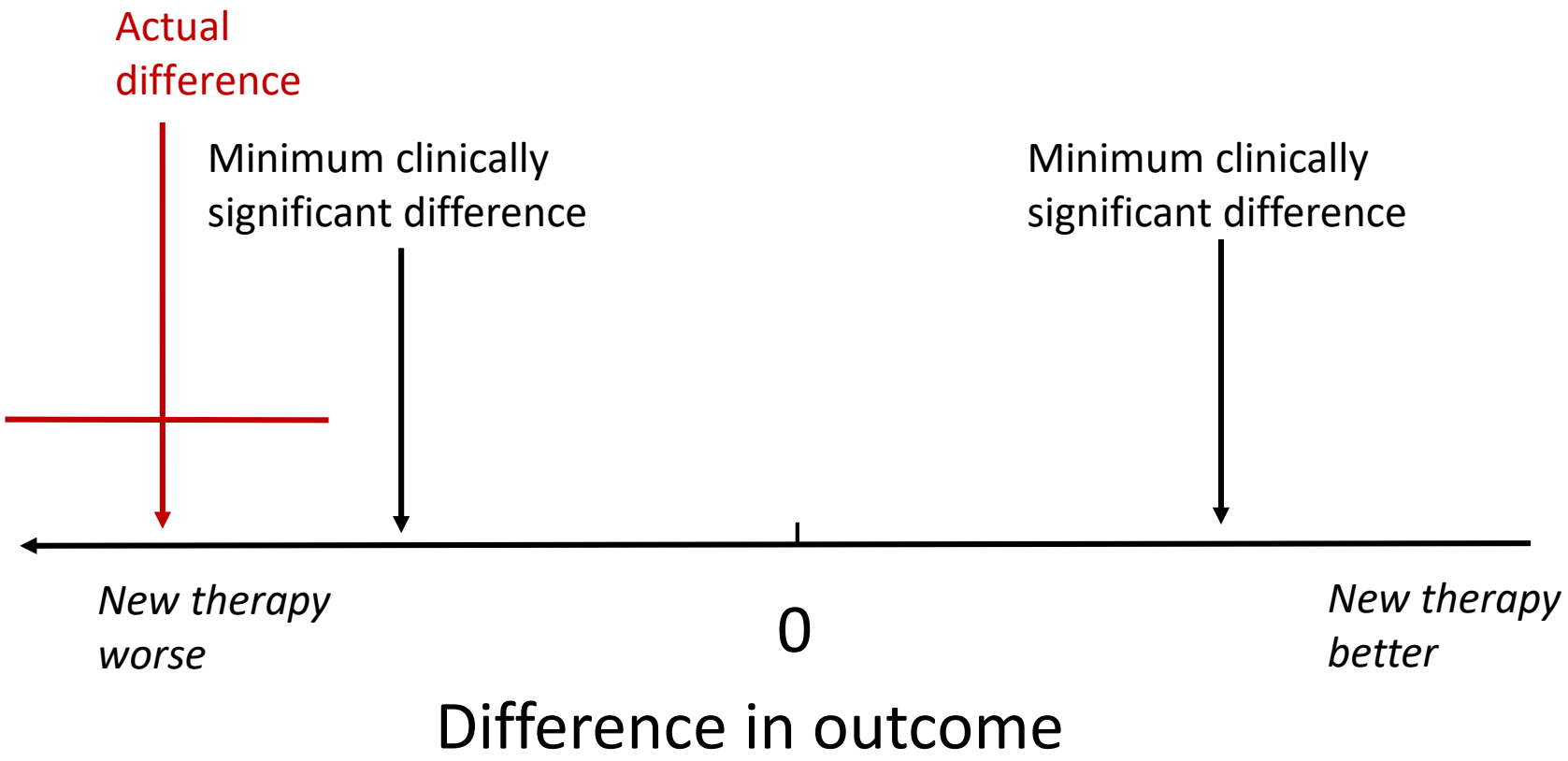
Statistically significant; Not better or worse; a tie; Non-inferior



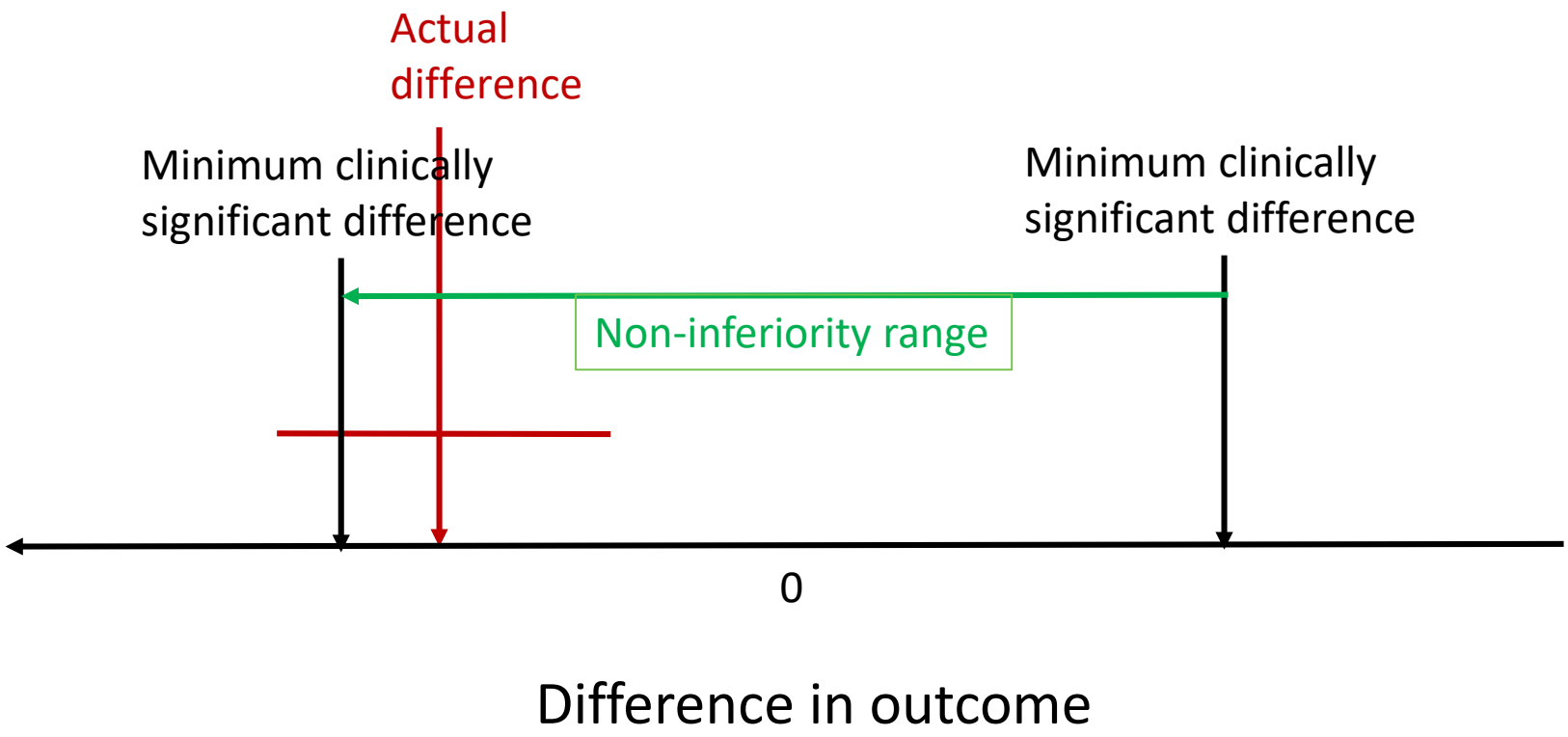
Statistically significant; Uninterpretable result for non-inferiority



Statistically significant; New therapy is inferior



- **If the 95% confidence interval for the difference overlaps the MCID threshold, the true difference could be on either side of the MCID even though the point estimate is within the range of non-inferiority.**
  - If the lower bound of the 95% CI is less than the MCID → cannot claim non-inferiority.



Statistical significance and interpretation in comparative effectiveness research

- When comparing two active treatments, the question is whether one treatment is enough better to prefer it.
- **Moral of the story**: deciding whether a treatment X is better than Y requires knowing how much better it would have to be to switch from Y to X.

Randomized trials

- Pragmatic trials
- Cluster trials
- Adaptive trials
- Non-inferiority trials