# Statistics of Clinical Trials

NEJM Resident Course

May 13, 2016

Reading a clinical trial paper

Non-inferiority trials

# Important questions to ask: design

- What was primary endpoint? Well chosen, clinically relevant?
- Was the design adequately powered for the clinically relevant endpoint
- Did the randomization produce balance?
- Did the analysis reflect the design?

# Important questions to ask: analysis

- Did the randomization produce balance?
- Should the primary analysis be intent to treat or per protocol.
- What does the primary analysis show?
  - Focus on confidence intervals, not p-values

# Important questions to ask: interpretation

- Was the rate in the control group close to the design specification?
- Are the trial results consistent? Across subgroups, across secondary endpoints?
- Do the conclusions in the discussion match those in the abstract

# Goals of NI trials

The NI design has one explicit goal and one implicit goal.

- Explicit goal: demonstrate that T is as effective, or nearly as effective, as best available therapy, C.
- Implicit goal: demonstrate that T is better than placebo no treatment (labeled P for placebo).

Ordinarily, both must be true for T to be a therapeutic option

- However, the FDA is required by law to approve drugs if they are shown to be "superior to no treatment".

# Goals of NI Designs

The ideal study design would be a three-arm design, with P, C, and T.

But a placebo arm is usually unethical

The non-inferiority trial provides a direct comparison of T to C, but it does not provide a direct comparison of T to P.

# Dealing with uncertainty

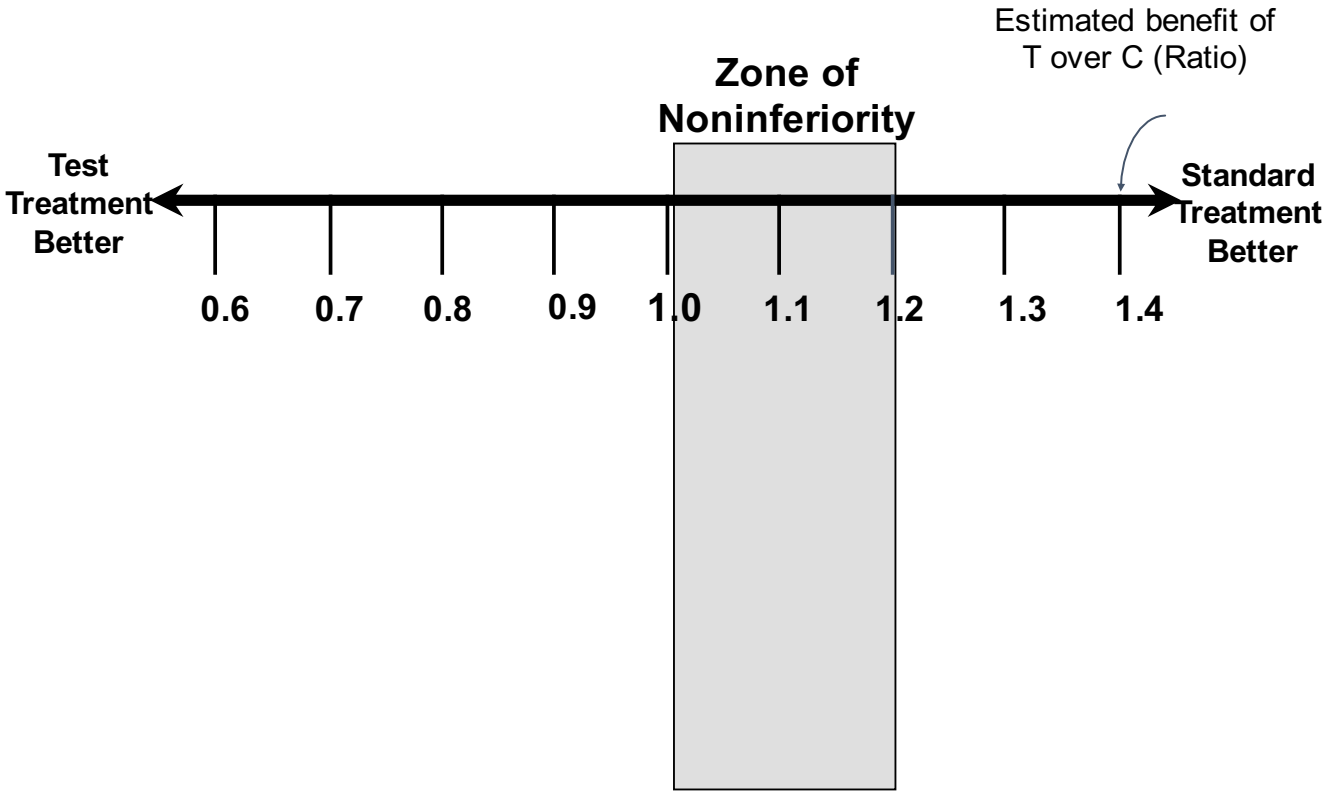
In 2010, FDA issued a Guidance recommending a two-step strategy for demonstrating non-inferiority:

1. Use the historical evidence to calculate an estimate 95% confidence interval for relative risk, control vs placebo.

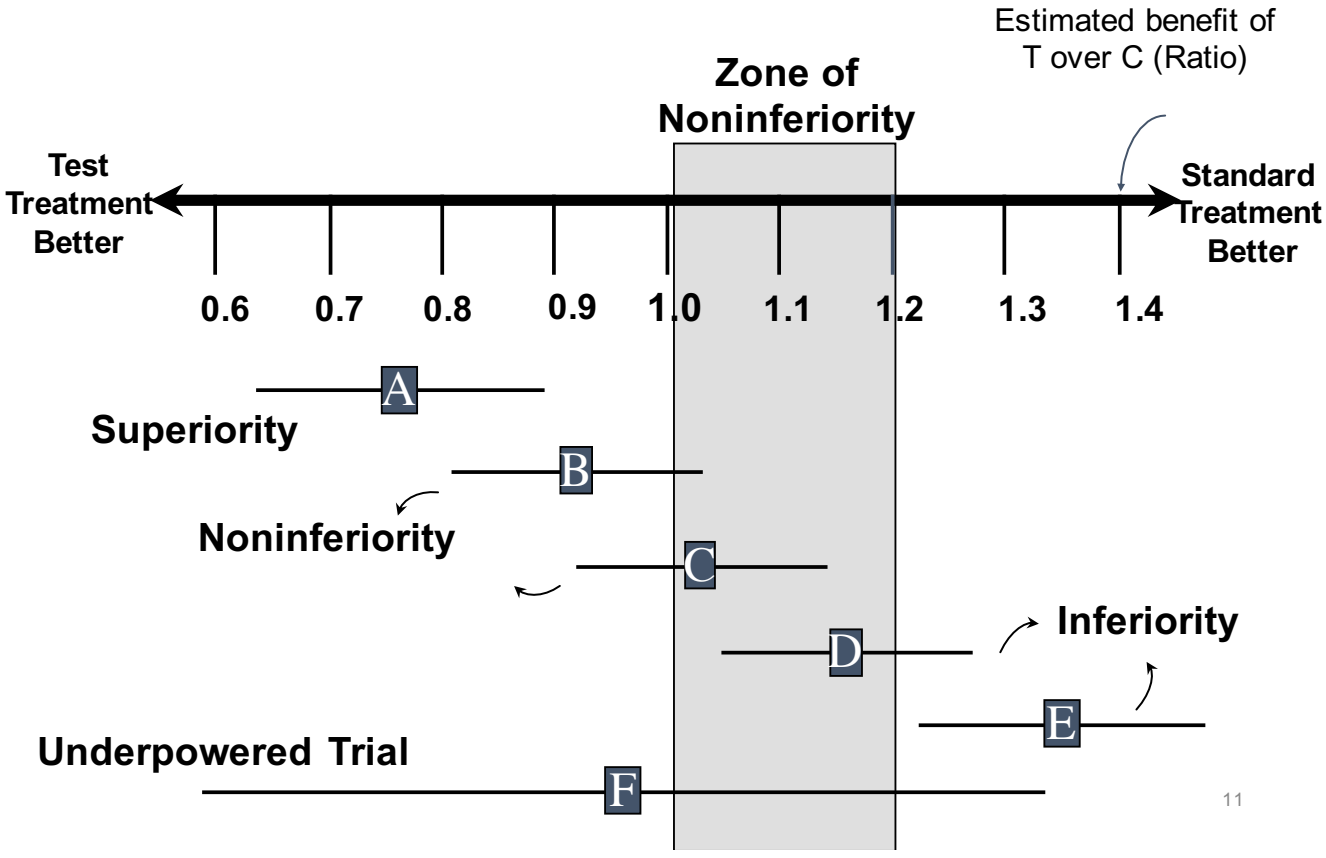
   - Set M equal to the lower limit of the 95% CI (the worst case estimate of the benefit of C relative to P)

2. Perform a NI trial to demonstrate that relative risk, Treatment vs Control. $<$ M

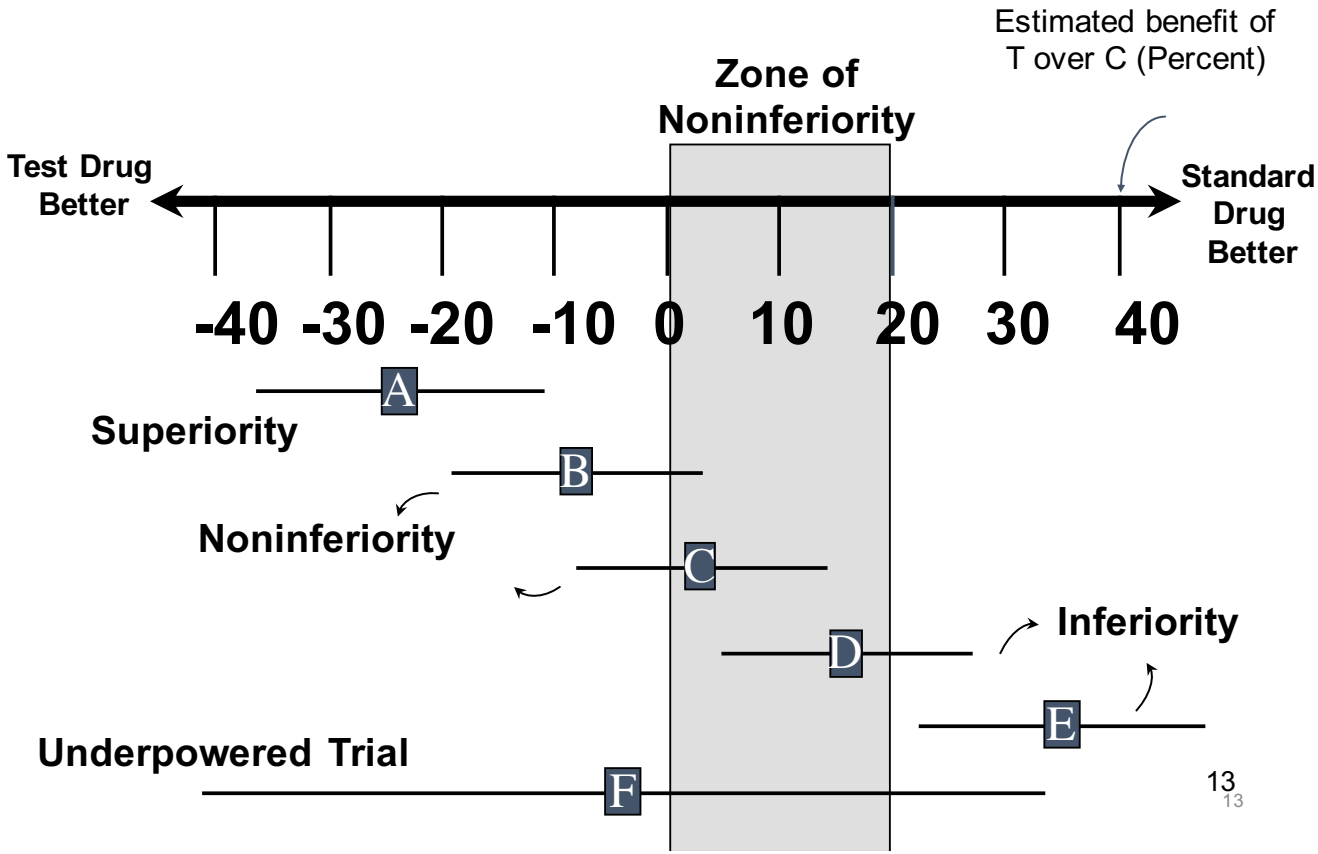
## Ratios of Event Rates :
## Test Drug/Standard Drug



## Ratios of Event Rates :
## Test Drug/Standard Drug



## Difference in Event Rates :
## Test Drug - Standard Drug



# Important issues in analysis of NI Trials

The role of null and alternative hypotheses are reversed in NI trials.

Suppose on relative risk scale, the hazard ratio of Treatment vs control should be no larger than 1.25

- The null hypothesis ($H_0$) is: Relative Risk $\geq 1.25$
- The alternative hypothesis ($H_A$) is: Relative Risk $< 1.25$

Small $p$-values lead to rejection of $H_0$

- Small p-values are not evidence of a treatment difference.

Sample size should provide adequate power (at least 80%) to reject $H_0$.

# Issues in the analysis of NI trials

Features of a trial which may lead to treatment differences appearing smaller may in appropriately lead to claim of NI.

- Crossovers, non-adherence, subsets of patients for whom T or C is not likely to be effective

Intent to Treat analysis (ITT) may be biased in NI trials.

Safest to provide both

- ITT analyses
- Per protocol (PP) analyses

# Threats to Validity of NI Designs

Non-inferiority designs have some intrinsic limitations that make them more difficult to design and more vulnerable to problems than the superiority design.

The following are the most important issues:

- Assay Sensitivity
- Assay Constancy
- Efficacy Creep

# Assay Sensitivity

A trial that demonstrates non-inferiority does not demonstrate efficacy. Both treatments could be effective or both could be ineffective in the setting of the trial.

The assumption that the comparator is effective in the current trial is known as assay sensitivity.

Many factors can contribute to assay insensitivity, e.g.,

- Lower than expected event rates
- Poor adherence
- Use of concomitant medications

# Assay Constancy

Assay constancy is the assumption that the size of the treatment effect is the same as in the historical trials

This is problematic because treatment of most diseases evolves rapidly.

# Efficacy Creep

Efficacy creep (sometimes called Biocreep): a slightly inferior treatment becomes the standard for the next generation of non-inferiority trials.

- If this happens repeatedly, one can end up with a standard that is no better than placebo.

To defend against efficacy creep, the active comparator should always be the "best" available therapy (BAT).

- . . . but the available evidence doesn't always provide a clear choice of the best treatment.

# Questions to ask about NI trials

- Is the claim of NI supported by a biological rationale?
- Might the effect of Active Control (vs placebo) have been different in current trial?
  - Changes in administration of agent,
  - Differences in populations using the drug or in endpoint determination
- Has long term follow-up changed the thinking of the value of the active control?
- Does the analysis use the best available historical data on active control to estimate both treatment effects and uncertainty in the estimate?

# Questions to ask about NI trials...

- Is an estimated NI margin clinically relevant? Was it specified in advance of the analysis?
- Is a reduced therapeutic effect for the test agent balanced by other benefits?
- What is the margin of error (confidence interval) in the estimate of possible loss of efficacy?
- Are results consistent across related endpoints?
- As in all trials, treatment effects measured in NI analyses are estimates of population effects, not predictions of efficacy for individuals Is there a clear signal to the treating clinician on when to use the active control vs the new treatment?

# Summary

Placebo-controlled trials are not always ethical. A new treatment might represent an advance even though it is no more efficacious than an existing treatment. Thus, NI trials are here to stay.

The International Conference on Harmonization guidelines say that

> *A suitable active comparator should be a widely used therapy whose efficacy has been clearly established and quantified in well-designed superiority trials and which can be expected to have similar efficacy in the contemplated trial.*

# Useful references

D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority Trials: Design concepts and issues – the encounters of academic consultants in statistics. *Stat Med* 2003;22:169-86.

Ellenberg SS, Temple R. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1. Ethical and scientific issues. *Ann Intern Med* 2000;133:455-63.

Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials, 5th ed.* 2015; Springer.