What's inferior about a non-inferiority trial?

Jorgen Seldrup
Centre for Statistics in Drug Development (CSDD)
Quintiles

Nothing!

The problem is:

It has been hijacked by the statisticians

Since clinicians cannot agree, let the statisticians solve the problem!

(Missing value problem)

Outline

- Design of NI trials
- Implementation of NI trials
- Analysis of NI trials
- Example

Focus on Design

Errors in analysis –
the trial may be rescued

Errors in design and the trial is 'dead'

Demonstrating Efficacy

(theory)

Choice of control group (ICH E10)

- Historical
- No treatment
- Placebo
- 'dose-response'
- Active treatment

CFR, 21 part 314.126: Adequate and well-controlled studies

Demonstrating Efficacy

(expectation)

- Superiority
  - Placebo controlled RCT
  - Active controlled RCT
- Non-inferiority
  - Active controlled RCT

Label

EFFICACY

Justification of NI Design

- Placebo is unethical
  - Accepted active treatment exists
  - Delaying or omitting treatment may be harmful
- Risk/Benefit profile
  - Fewer/less serious AEs
  - Treatment administration advantages

Simple Statistical Setting

Hypotheses for a NI trial (symbolic):

H_0: $T \leq C - \delta$

H_1: $T > C - \delta$

(T = Test/new, C = Control/active, δ = NI margin)

Proportions/Ratios/Means

Simple Statistical Setting

Hypotheses for a NI trial (symbolic):

H_0: $T \leq C - \delta$

H_1: $T > C - \delta$

(T = Test/new, C = Control/active, δ = NI margin)

Proportions/Ratios/Means

Hot Design Topics

- NI margin
  - Clinical margin vs statistical margin
  - Clinical acceptability
  - Statistical reasoning
    - Constancy assumption
    - Assay sensitivity (implementation topic)
    - Putative placebo analysis
- "Biocreep"
- Switching

Clinical Margin

- Clinical acceptability
  - Conceptually ‘unfamiliar’
  - Subjective
    - Agreement
      - Investigator/Sponsor/Regulator

Constancy Assumption

The effect of the active control treatment does not change from the past to the present

- Generally cannot be verified
  - Patient population
  - Background therapy
  - Medical practice
  - Endpoints

… so, what next?

Assay Sensitivity

"… the active control would have been superior to placebo in the NI trial had a placebo been used …"

… but it wasn't!

… so, what next?

Next

We statisticians do our best

(but is it good enough)

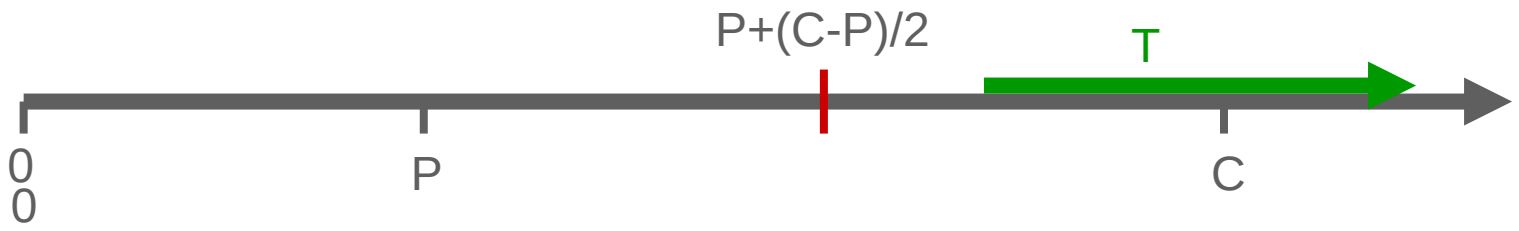
… an approximate answer to the right question is better than the right answer to the wrong question …

Statistical Margin

Premise:

On existing evidence:

- The effect of the new treatment should be ‘close’ to that of the active control treatment
- The new treatment should retain a certain proportion of the effect of the active control treatment (e.g., 50%)



Putative Placebo Analysis

(Hasselblad and Kong, DIJ:2001)

$$T/P = T/C \times C/P,$$

e.g., T/C and C/P are relative risks

NOTE:
T/C is from the present NI trial
C/P is from a meta-analysis of historical trials

Relies largely on meta-analysis!
… but who believes in meta-analysis?

T vs P

- Need to estimate (or test) the difference ‘T-P’ (or some fraction)
- Need to account for within- and across trial variability
  - CBER fixed margin method
  - Synthesis method
- Pragmatic approach
  - ‘smallest lower bound’
  - Lower confidence limit from meta-analysis

"Biocreep"

… occurs when an inferior test treatment becomes the active control treatment in the next NI trial and so on until the active control is no better than placebo

Cure:

Always use the 'best'/same comparator/ active control treatment

Switching – NI to Sup

(non-inferiority trial) … If the lower limit of the calculated 95% confidence interval (CI; 2-sided) for the difference T - S in overall success rate, is greater than -15%, T will be considered at least as effective as S. In accordance with the PtC (Ref), should the lower limit of the 2-sided 95% CI be greater than 0 (zero), the p-value associated with an appropriate test of superiority will be calculated. A p-value of < 0.01 will be considered convincing evidence of the superiority of T…

PtC on switching between superiority and non-inferiority (CPMP/EWP/482/99, London, 27 July 2000)

IV.1 (PtC): "In this case (*sic*, i.e., switching) it is acceptable to calculate the p-value associated with a test of superiority and to evaluate whether this is sufficiently small to reject convincingly the hypothesis of no difference"

Balancing Benefit/Risk

NI efficacy trial – the hidden agenda

- If NI design is justified by 'maybe not quite as efficacious but fewer AE'
  - Dual primary objective: NI for efficacy; superiority for safety
  - No need for adjustment (but the interpretation will only be 'success' if both objectives are achieved)
  - Sample size considerations

Implementation/conduct

Adequate and well-controlled studies

See CFR Title 21, §314.126:
Adequate and well-controlled studies

(Not so) Hot Analysis Topics

- Analysis populations
  - E9: "… in a NI trial use of the FAS/ITT is generally not conservative and its role should be considered very carefully"
  - PtC Switching: "…in a NI trial the FAS and the PP analysis set have equal importance … and should lead to similar conclusions"
- Secondary analyses (?NI or superiority)

Example

Indication:

Invasive fungal disease; serious and life-threatening

Background to selection of NI margin:

P = 0%

S (active) = 37% (range 14%-83%)

Recently one adequate and well-controlled NI trial (margin 20%) had shown

C = 53%

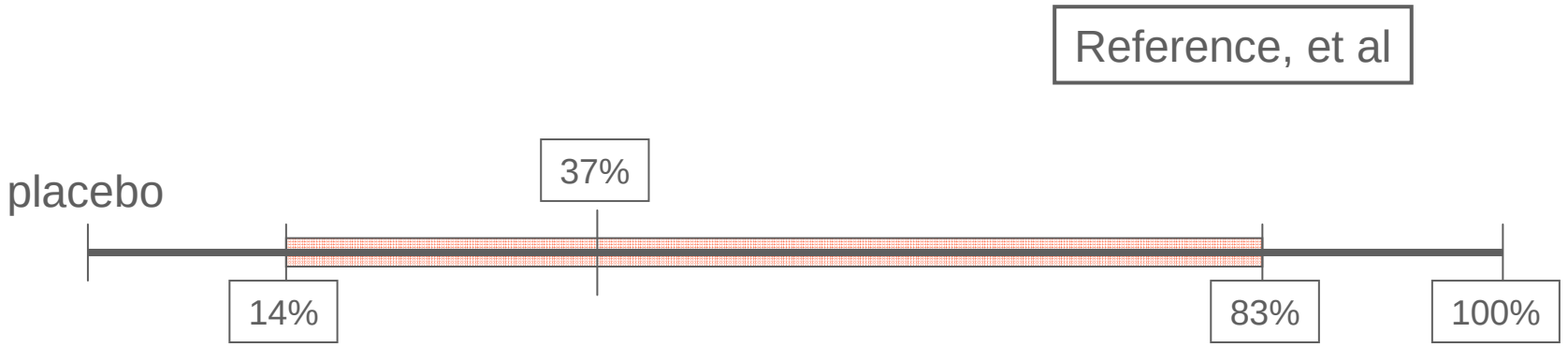
to be superior to

S = 32%

S

QUINTILES TRANSNATIONAL

Successful outcome rates

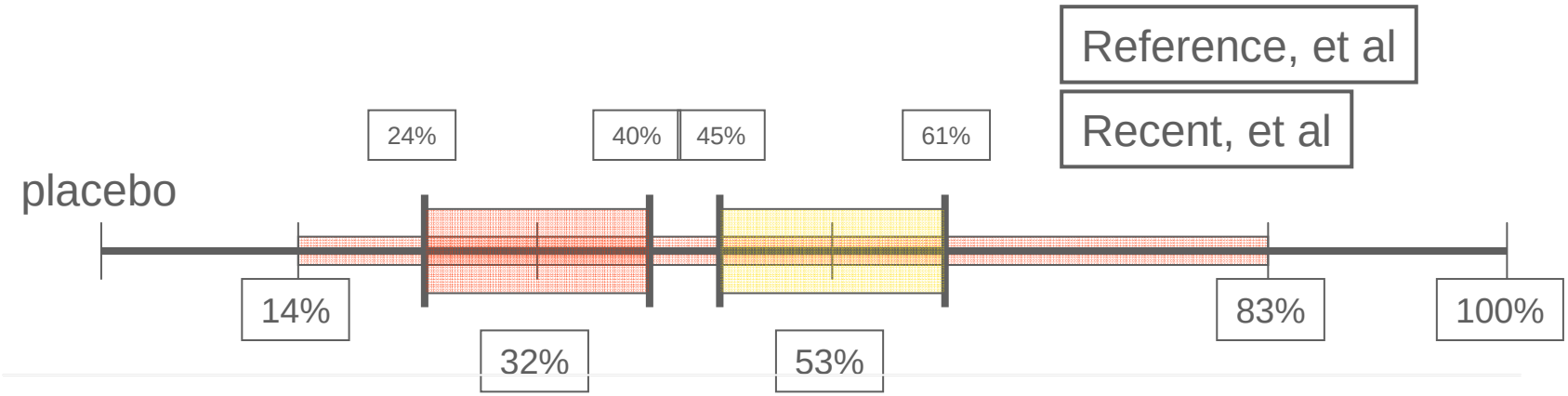


S

C

QUINTILES TRANSNATIONAL

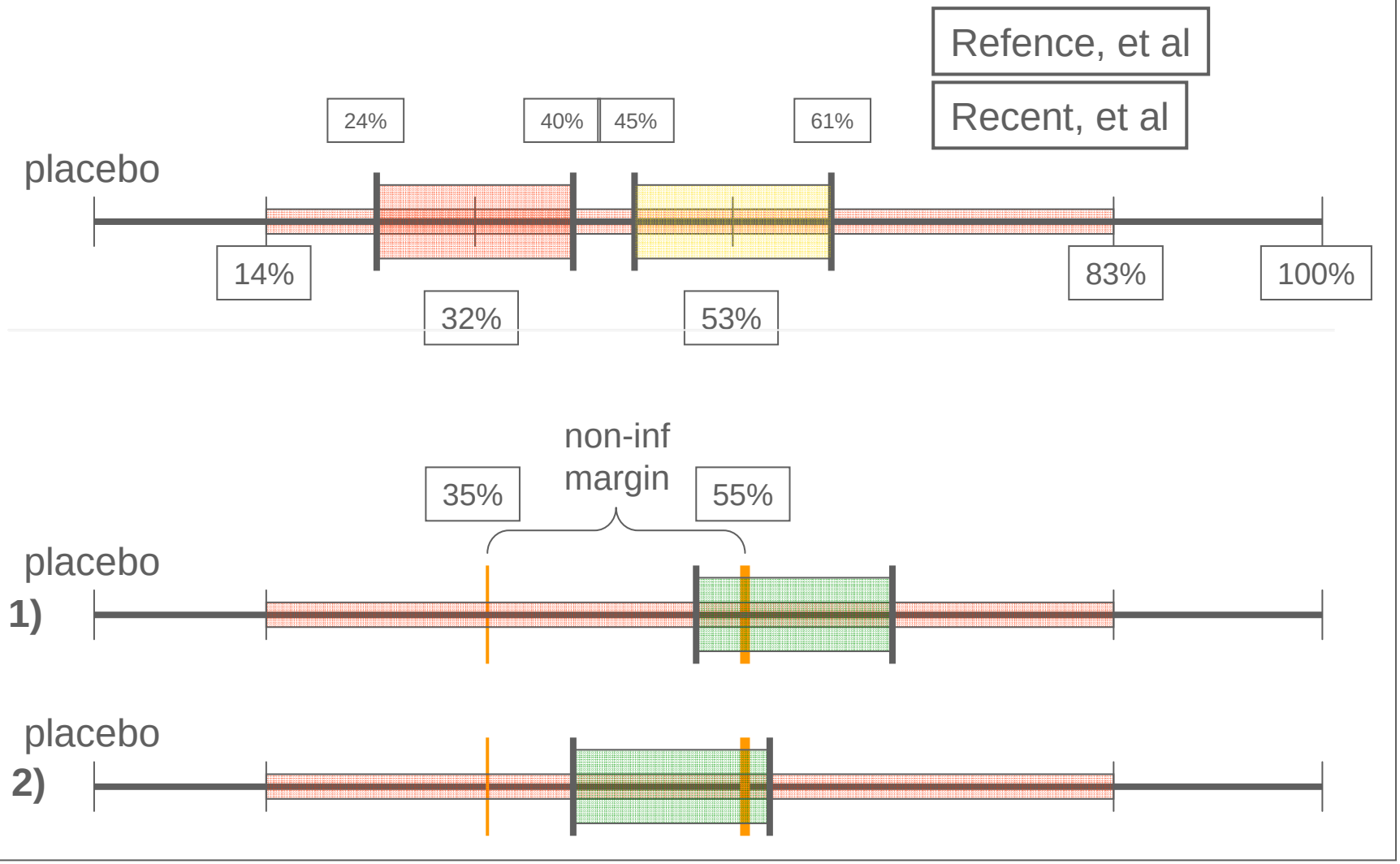
Successful outcome rates



S C T

QUINTILES TRANSNATIONAL

Successful outcome rates



Protocol statement

The proposed margin of 20% will thus allow to demonstrate activity of T in the clinically expected and acceptable range, i.e. superiority over no treatment, and efficacy at least as good as S. Non-inferiority to C at the 20% level would indicate that T may offer certain clinical benefits not available with currently approved drugs of the same class.

The following hypothesis will be tested, where R is the successful outcome rate:
H_0: $R(T) \leq R(C) - \delta$
H_1: $R(T) > R(C) - \delta$.
A sample size of 125 patients per treatment group is calculated based on the following assumptions (NQuery Advisor, v5, module PTE1a):
Overall outcome rate R = 55%
δ = 20%
Power = 90%
Significance level = 0.025 (1-sided)
Allowing for 30% of patients not completing the trial satisfactorily (i.e. unevaluable for overall outcome at Day 42), approximately 180 patients per group or 360 patients will be enrolled.

Conclusion/Position

- NI trial is OK for demonstrating efficacy (when used appropriately)
- Have a meaningful pragmatic discussion about the choice of δ (sample size)
- Don’t get hung up on problems we cannot solve (placebo *is* missing – don’t invent it)
- Plan on an “adequate and well-controlled study”