Non-inferiority trials and switch from non-inferiority to superiority

D Costagliola

U 943 INSERM and UPMC Paris 06

References

- ICH E 9 et E10
- Points to consider on biostatistical methodological issues from recent CPMP discussions on licensing applications: superiority, non-inferiority and equivalence

Clinical trial objectives

- Trials comparing a new treatment (or a strategy) to a reference treatment
  - Showing the superiority of the new treatment
    - Is N better than R ?
      - Pre-treated patients
  - Showing the non-inferiority of the new treatment
    - Is N doing not worse than R ?
      - Naive patients
  - Showing the equivalence of the new treatment
    - Is N doing as well as (neither better not worse) R ?
      - Bio-equivalence (different formulation of the same drug)

Example - 1

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Tenofovir DF, Emtricitabine, and Efavirenz vs. Zidovudine, Lamivudine, and Efavirenz for HIV

Joel E. Gallant, M.D., M.P.H., Edwin DeJesus, M.D., José R. Arribas, M.D., Anton L. Pozniak, M.D., Brian Gazzard, M.D., Rafael E. Campo, M.D., Biao Lu, Ph.D., Damian McColl, Ph.D., Steven Chuck, M.D., Jeffrey Enejosa, M.D., John J. Toole, M.D., Ph.D., and Andrew K. Cheng, M.D., Ph.D., for the Study 934 Group*

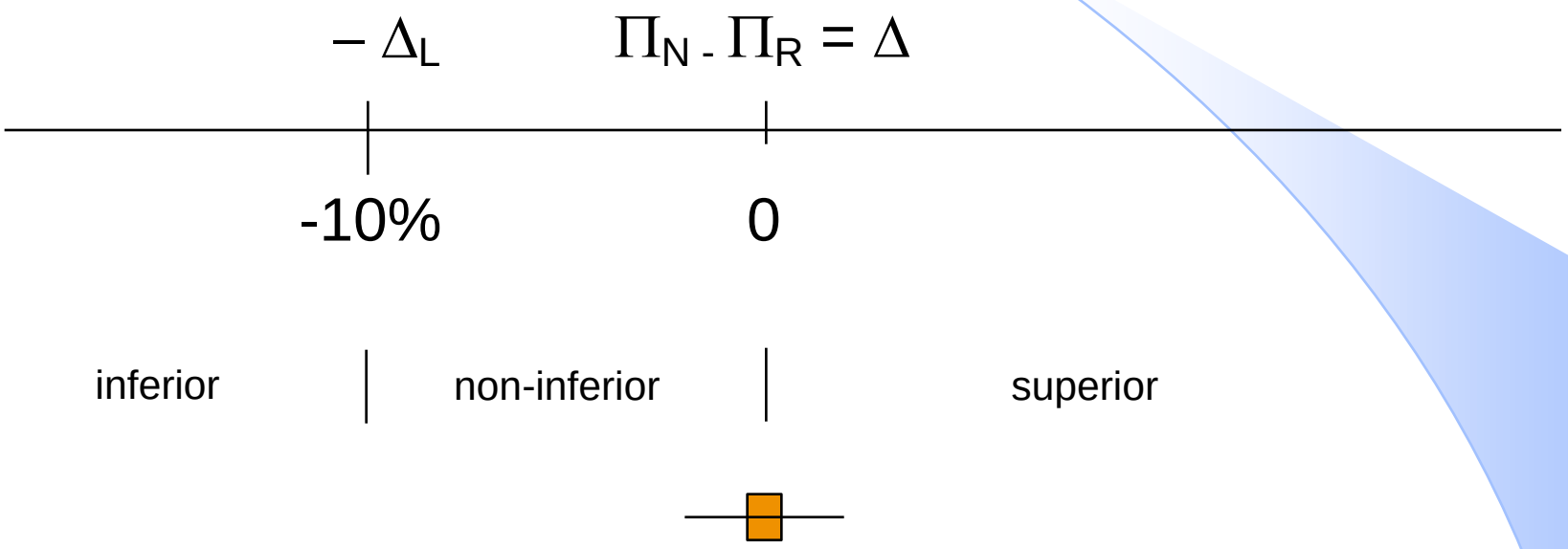
Example - 2

EFFICACY ANALYSIS

The primary objective was to assess the noninferiority of the regimen of tenofovir DF, emtricitabine, and efavirenz to the regimen of zidovudine, lamivudine, and efavirenz as measured by HIV RNA levels of less than 400 copies per milliliter through week 48, defined according to the algorithm of the Food and Drug Administration (FDA) for the time to loss of virologic response, which requires confirmation (two consecutive values) of response or of no response (missing data or early termination of participation in the study was considered to be failure).[5] The 487 eligible patients without baseline resistance to efavirenz who underwent randomization and received treatment were the predefined population for the primary end-point analysis. The secondary objective was to assess the noninferiority of tenofovir DF, emtricitabine, and efavirenz to zidovudine, lamivudine, and efavirenz as assessed by HIV RNA levels of less than 50 copies per milliliter and changes in the CD4 cell count.

Definition of non-inferiority

- N is not doing worse than R



Choice of the non-inferiority limit - 1

- Clinical decision, not statistical
- The largest difference clinically acceptable
- < = difference used in superiority trials of the same domain
- To warrant that the new product is doing better than placebo in trials with no placebo

A working case in diabetes: HbA1c the risk of death - 1

- In diabetes, for new drugs the most common endpoint is HbA1C
  - Non inferiority margin usually taken as 0.6 %
  - Superiority trials usually try to demonstrate a 1% difference

A working case in diabetes: HbA1c the risk of death - 2

- Each 1% reduction in updated mean HbA1c was associated with reductions in risk of
  - 21% for any end point related to diabetes (95% confidence interval 17% to 24%),
  - 21% for deaths related to diabetes (15% to 27%),
  - 14% for myocardial infarction (8% to 21%), and
  - 37% for microvascular complications (33% to 41%).
  - No threshold of risk was observed for any end point.

Stratton IM et al. UKPDS 35. BMJ 2000;321:405-412

A working case in diabetes: HbA1c the risk of death - 3

- Is it possible to define a non-inferiority limit clinically acceptable in this context?

Choice of the non-inferiority limit - 2

- As defining a non-inferiority limit implies to accept some loss
  - There must be some advantage to use the new product
    - easyness
    - safety
    - costs
    - …

Example - 3

STATISTICAL ANALYSIS

The regimen of tenofovir DF, emtricitabine, and efavirenz was to be considered not inferior to the regimen of zidovudine, lamivudine, and efavirenz if the lower bound of the 95 percent confidence interval for the difference between the two groups, those receiving tenofovir DF, emtricitabine, and efavirenz (the tenofovir–emtricitabine group) minus those receiving zidovudine, lamivudine, and efavirenz (the zidovudine–lamivudine group) for the primary end point (in the proportion of patients with an HIV RNA level of less than 400 copies per milliliter) was no lower than –13 percent. Assuming a response rate of 70 percent at week 48 for the zidovudine, lamivudine, and efavirenz regimen and a one-sided type I error of 2.5 percent, the planned sample size of 500 patients provided the study with 85 percent power to demonstrate the noninferiority of the tenofovir DF, emtricitabine, and efavirenz regimen. Substitu-

- Non-inferiority limit 13% (between 7 and 15%)
- Expected success rate 70% (EFV vs IND)
- One-sided Type I error 2.5%
- Power 85 %

Other issues

- Internal validity
  - Limited
    - protocol deviation,
    - lack of adherence,
    - lost to follow-up,
    - and missing data
  - Because they biased the result towards no difference
- External validity
  - Choice of the reference treatment
    - Known efficacy
    - Placebo group when possible
  - Study population
    - Same as the one in which the reference treatment was shown efficacious
  - Endpoint(s)
    - Same as the one(s) used to show the reference treatment efficacy
  - Expected efficacy from the reference treatment observed in the current trial

Comparison test

- Superiority (two-sided)
  - $H_0 : \Pi_N = \Pi_R$
  - $H_1 : \Pi_N \neq \Pi_R$
- Superiority (one-sided)
  - $H_0 : \Pi_N = \Pi_R$
  - $H_1 : \Pi_N >$ à Π_R
- Non inferiority (one-sided)
  - H'_0: $(\Pi_N - \Pi_R) = \Delta < -\Delta_L$ (N is inferior to R)
  - H'_1: $(\Pi_N - \Pi_R) = \Delta >= -\Delta_L$ (N is non inferior or superior R)
  - The non inferiority limit Δ_L influences the result of the analysis

Sample size 1

- Δ_L is usually smaller than the interesting difference in a superiority trial in the same field
  - The sample size tends to be larger

Sample size 2

Table 2. Sample sizes per arm for noninferiority trials, by power, delta and expected response rate in the control arm; the efficacy of the new drug is assumed to be equivalent for the purposes of calculating sample sizes.

Expected response rate in control arm	Delta 12% 80% power	90% power	Delta 10% 80% power	90% power
50%	273	365	393	526
55%	270	362	389	521
60%	262	351	377	505
65%	249	333	358	479
70%	229	307	330	442
75%	205	274	295	395
80%	175	234	252	337
85%	139	187	201	268
90%	99	132	142	190

Hill A AIDS 2008;22:913-921

Analysis plan

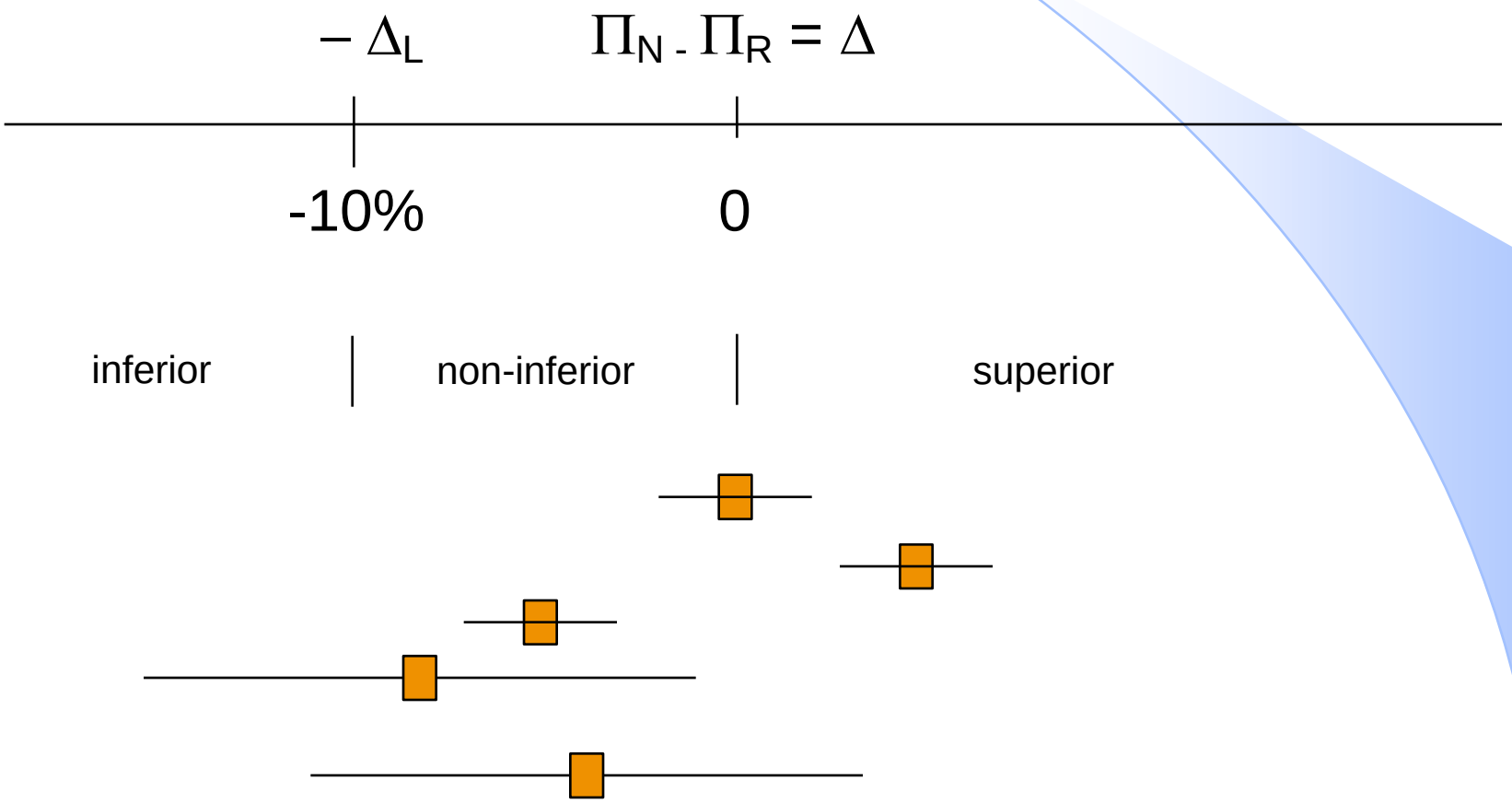
- Results
  - Confidence intervals of the difference
  - More rarely a p-value
- Both ITT and per protocol analyses should be conducted and give the same results
- Analysis of compliance to treatment and protocol deviation (+++)

The conclusion is based on

- The lower limit of the confidence interval of the estimated difference compared with the non inferiority limit Δ_L

Definition of non-inferiority

- N is not doing worse than R



Exemple - 4

RESPONSE TO TREATMENT

At week 48, 206 of the 244 patients (84 percent) in the tenofovir–emtricitabine group and 177 of the 243 patients in the zidovudine–lamivudine group (73 percent) reached and maintained HIV RNA levels of less than 400 copies per milliliter, which was the primary end point (Fig. 2). The 95 percent confidence interval for the difference between the two groups was 4 to 19 percent (P=0.002), which excludes the inferiority of the tenofovir DF, emtricitabine, and efavirenz regimen. The confidence interval for the difference also excludes zero, indicating a significantly greater response with the tenofovir DF, emtricitabine, and efavirenz regimen. At week 48, 194 of 244 patients (80 percent) in the tenofovir–emtricitabine group and 171 of 243 patients in the zidovudine–lamivudine group (70 percent) reached and maintained HIV RNA levels of less than 50 copies per milliliter. The 95 percent confidence interval for the difference between the two groups was 2 to 17 percent (P=0.02), which excludes the inferiority of the tenofovir DF, emtricitabine, and efavirenz regimen and indicates a significantly greater response with this regimen. Similar statistically significant differences were observed in the intention-to-treat population (509 patients) on the basis of HIV RNA levels of less than 400 copies per milliliter (81 percent in the tenofovir–emtricitabine group vs. 70 percent in the zidovudine–lamivudine group; 95 percent confidence interval for the difference, 3 to 18 percent; P=0.005) or HIV RNA levels of less than 50 copies per milliliter (77 percent vs. 68 percent, respectively; 95 percent confidence interval for the difference, 1 to 16 percent; P=0.03).

PP : 84% vs 73%,
95% CI : 4-19, N = 487
ITT : 81% vs 70%,
95% CI : 3-18, N = 509

Exemple - 5

In this large, randomized trial, the tenofovir DF, emtricitabine, and efavirenz regimen fulfilled the criteria for noninferiority to the zidovudine, lamivudine, and efavirenz regimen. The results also indicate significantly greater responses to the tenofovir DF, emtricitabine, and efavirenz regimen (as defined by the FDA's algorithm for the time to loss of viral response) as compared with the well-established regimen of zidovudine, lamivudine, and efavirenz.[8,9] The regimens also differed in their effect on immune reconstitution: in the tenofovir–emtricitabine group, there was a significantly greater increase in the total CD4 cell counts and the CD4 percentages.

Interpreting a non-inferiority trial as a superiority trial

- No majors issues, but is the difference of clinical significance ?
  - Depending on
    - The reference treatment
    - The power
    - The effect size
    - The analysed population
    - The trial quality
  - The p value for the superiority test is derived from the ITT analysis

Example - 6

RESULTS

Through week 48, significantly more patients in the tenofovir–emtricitabine group reached and maintained the primary end point of less than 400 copies of HIV RNA per milliliter than did those in the zidovudine–lamivudine group (84 percent vs. 73 percent, respectively; 95 percent confidence interval for the difference, 4 to 19 percent; P=0.002). This difference excludes the inferiority of the tenofovir DF, emtricitabine, and efavirenz regimen, indicating a significantly greater response with this regimen. Significant differences were also seen in the proportion of patients with

The right p value for the ITT analysis was p=0.005 (81 % vs 70 %)

Example - 7

- What would the gain be?
  - Not listed in the conclusion
  - Difference on the loss of leg fat at W48?
  - Change in LDL-cholesterol?
- No longer an issue if there is a clinically significant superiority

Conclusion

- If one accepts a loss of chance, what is the expected gain?
- The choice of the non-inferiority limit is critical
  - It is a clinical, not a statistical decision
  - Should warrant that the new product is better than placebo
  - Typically 7-12% in the recent trials in HIV
- The ITT analysis is no longer the main analysis
  - Both ITT and per protocol are important
  - The difference in the number of patients included in each analysis is an indicator of the study quality
- No major issues in switching from non-inferiority to superiority