Non-inferiority Trials – Hypotheses and Analyses

Gang Chen[1], Yongcheng Wang[2], George Chi[1], Kevin Liu[1]

[1] Clinical Biostatistics, Global Drug Development, J&J PRD,
[2]Food and Drug Administration

November 1, 2004, BASS XI, Savannah, Georgia

Outline

- Non-inferiority (NI) hypotheses
  - fixed margin
  - fraction retention
- Analysis methods
- Example
- Major issues and summary

NI Hypotheses

Fraction retention/ Fixed margin

Notations

- Endpoint: time to event (e.g., survival, TTP)
- Hazard ratio: HR(T/C) and HR(P/C)
- Treatment effect: θ_1 = HR(T/C) -1
- Control effect: θ_2 = HR(P/C) -1
- Fraction retention of control effect:

  $\delta = 1 - \{\theta_1 / \theta_2\}$, or
- Fraction loss of control effect

  $1 - \delta = \theta_1 / \theta_2$,

  where, T, C and P are treatment, control and placebo respectively.

NI hypotheses – Fraction retention

- Fraction retention NI hypotheses:

$H_0: \theta_1/\theta_2 \geq 1 - \delta_0$ vs. $H_a: \theta_1/\theta_2 < 1- \delta_0$, or,

if $\theta_2 > 0$,

$H_0: \theta_1 - (1- \delta_0)\,\theta_2 \geq 0$ vs. $H_a: \theta_1 - (1 - \delta_0)\,\theta_2 < 0$.

NI Hypotheses-fraction retention
Selection of fraction retention

The selection of fraction retention depends on several factors:

- objective of active control trial
  - claim non-inferiority or equivalence
  - claim efficacy
- clinical judgment
- statistical judgment
  - distributional properties of the ratio of treatment effect vs. active control effect
  - mean effect size of active control
  - variability of active control effect

NI hypotheses – Fixed margin

- If fix control effect $\theta_2 = M_1 > 0$, and define margin $M = M_1 * \delta_0$, where δ_0 is a fixed level of fraction retention, then NI hypotheses become:

$$H_0: \theta_1/M_1 \geq \delta_0 \text{ vs. } H_a: \theta_1/M_1 < \delta_0, \text{ or}$$

$$H_0: HR(T/C) \geq 1+M \text{ vs. } H_a: HR(T/C) < 1+M$$

NI hypotheses-Fixed margin

Margin selection:

- Arbitrary margin: questionable
- Margin based on control effect ~ two CI method:
  Based on the lower limit (LL) of γ% CI for HR(P/C), i.e.

  Margin = δ_0*(LL γ %CI for HR(P/C) -1)

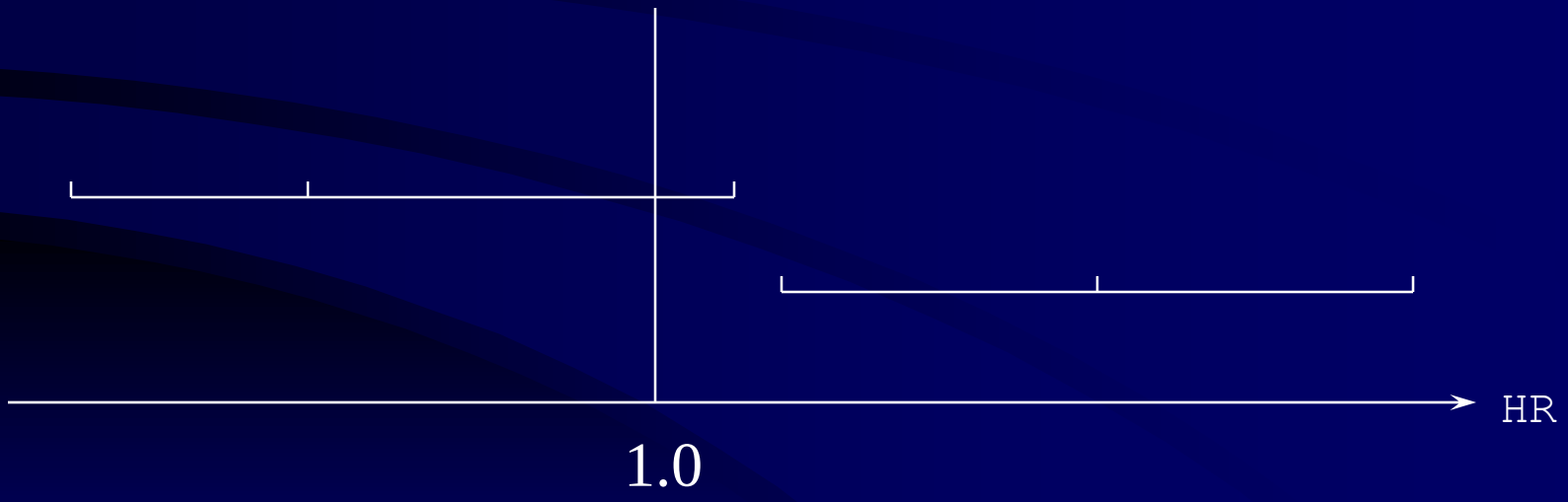
  e.g., δ_0 = .5 & LL of γ %CI = 1.2, then margin = .1

  If the 95% CI for HR(T/C) lies entirely beneath 1 + margin (NI cutoff), “non-inferiority” is concluded

NI hypotheses-Fixed margin

Two CI approach:

95 % CI for HR(T/C) γ % CI (cutoff) for HR(P/C)



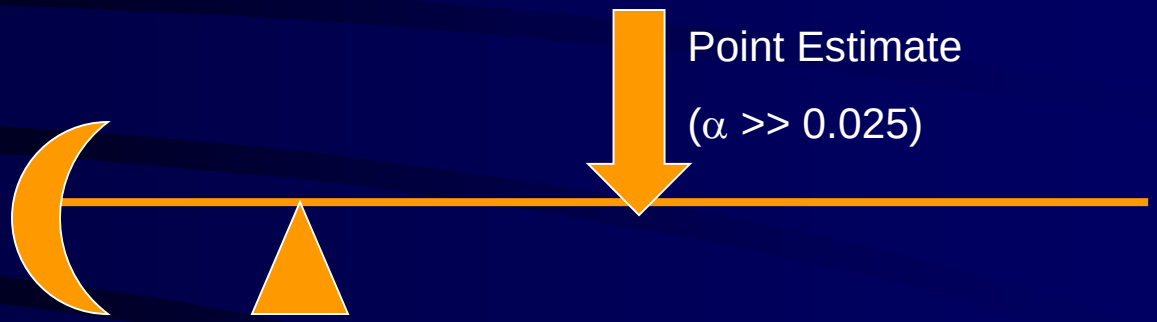
NI hypotheses-Fixed margin

Margin selection, for example:

- γ=0: margin = point estimate
- γ = .3: margin = LL of 30% CI
- γ=.95: margin = LL of 95% CI

NI hypotheses-Fixed margin

margin and type I error:



Assessment of control effect

- There should be some historical randomized, double-blind and placebo controlled studies involving the active control.
- Modeling active control effect using a meta-analysis (either random effects or fixed effects model).
  - Random effects model may be preferred because it provides a more appropriate standard error.
  - When there is only one or two historical active control trials, it is difficult to assess the control effect and the between study variability may not be appropriately assessed.

Assessment of control effect

- Constancy of the control effect: Current active control effect needs to be assessed with the following consideration
  - Changes in populations?
  - Changes in standard care, or medical practice (including concomitant medications)?
- Appropriate adjustment may be necessary if the constancy assumption my be wrong:
  - Adjustment for control effect size
  - Adjustment for characteristics of patient population

Interpretation of NI hypotheses

- The discussion and interpretation of fixed margin NI hypotheses and fraction retention NI hypotheses are given in [1] [2].

[1] George YH Chi, Gang Chen, Mark Rothmann, Ning Li (2003), Active Control Trials. Encyclopedia of Biopharmaceutical Statistics: Second Edition.

[2] Mark Rothmann, Ning Li, Gang Chen, George Y.H. Chi, Hsiao-Hui Tsou, and Robert Temple (2003), Design and analysis of non-inferiority mortality trials in oncology, Statistics in Medicine. Vol. 22: 239-264.

Statistical Tests

NI test procedure

Non-inferiority test procedure:

- Step 1: assessing control effect θ_2 based on historical randomized trials. If control effect is positive, then
- Step 2: assuming $\theta_2 > 0$ (control is effective) and formulate fraction retention NI hypotheses (or fixed margin hypotheses with $\theta_2 = M$):

  $H_0: \theta_1/\theta_2 \geq 1 - \delta_0$ vs. $H_a: \theta_1/\theta_2 < 1- \delta_0$, or, if $\theta_2 > 0$,

  $H_0: \theta_1 – (1- \delta_0)\, \theta_2 \geq 0$ vs. $H_a: \theta_1 – (1 - \delta_0)\, \theta_2 < 0$.
- Step 3: drawing inference with alpha < 0.05 for NI hypotheses and claiming NI.

NI test procedure

- One concern on NI test procedure: The false positive rate associated with the non-inferiority test procedure may be inflated. The details have been discussed in [1].

[1] Gang Chen, Yong-Cheng Wang, George Chi (2004), Hypotheses and type I error in active control non-inferiority trials, Journal of Biopharmaceutical Statistics, Journal of Biopharmaceutical Statistics. JBS, Vol. 14, No. 2, pp 301-313.

Statistical Tests

- Linear test (Rothmann)
- Ratio test (Wang)
- Two 95% CI
- CI for the ratio (H/K)
- Bayesian (Simon)

Linear test

- NI hypotheses: Assuming HR(P/C) > 1

$H_0^{(1)}$: $\log HR(T/C) \geq (1-\delta_0)\log HR(P/C)$

vs. $H_a^{(1)}$: $\log HR(T/C) < (1-\delta_0)\log HR(P/C)$

Linear test

- Test statistic for $H_0^{(1)}$ vs. $H_a^{(1)}$:

$$Z_{(1)}^{*} = \frac{\log\hat{HR}(T/C) - (1-\delta_0)\log\hat{HR}(P/C)}{\sqrt{s_1^2 + (1-\delta_0)^2 s_2^2}}$$

where $\hat{HR}(T/C)$ and $\hat{HR}(P/C)$ are the estimates of hazard ratios, and

$s_1 = s.e.(\log\hat{HR}(T/C))$, $s_2 = s.e.(\log\hat{HR}(P/C))$

Linear test
Normality, Power and Sample size

- Details given in the paper:

Mark Rothmann, Ning Li, Gang Chen, George Y.H. Chi, Hsiao-Hui Tsou, and Robert Temple (2003), Design and analysis of non-inferiority mortality trials in oncology, Statistics in Medicine. Vol. 22: 239-264.

Ratio Test

- Hypothesis:

$$H_0: \delta < \delta_0 \quad \text{vs.} \quad H_a: \delta > \delta_0$$

Ratio Test

- Estimate of δ:

$$\hat{\delta} = \frac{[\hat{HR}(P/C)-1]-[\hat{HR}(T/C)-1]}{\hat{HR}(P/C)-1} = 1 - \frac{\hat{HR}(T/C)-1}{\hat{HR}(P/C)-1}$$

where $\hat{HR}(T/C)$ and $\hat{HR}(P/C)$ are estimates of hazard ratios.

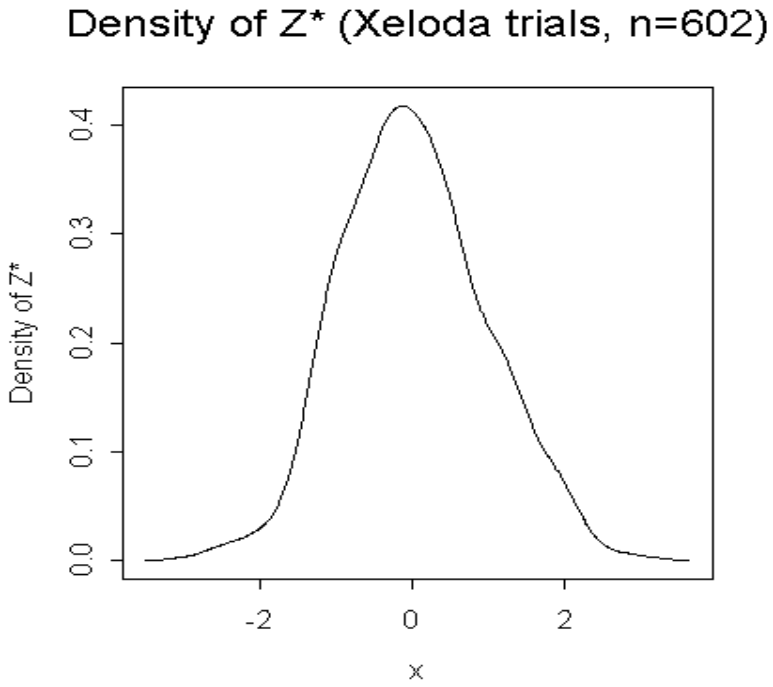
Ratio Test

- Test statistic:

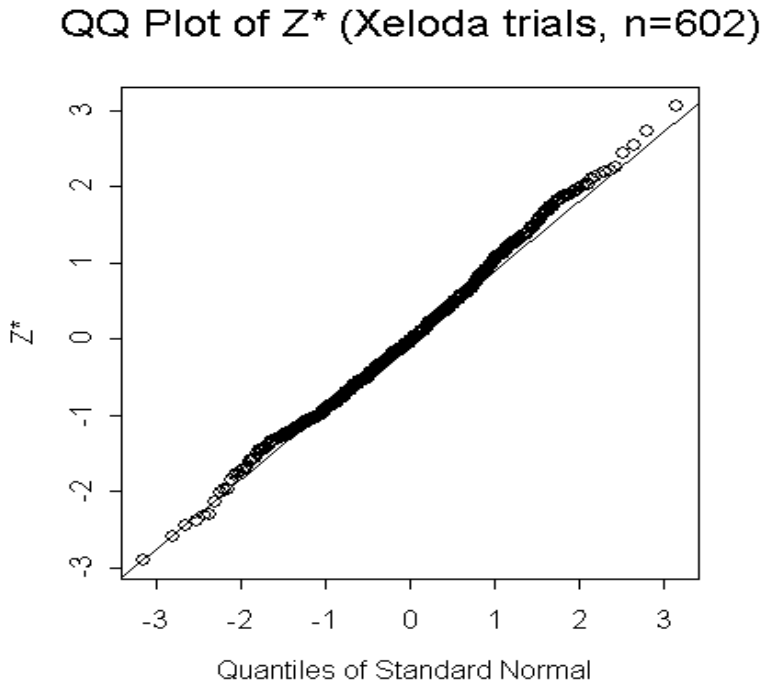
$$Z^* = \frac{\hat{\delta} - \delta_0}{\text{s.e.}(\hat{\delta})}$$

Concern: Is Z* normal?

Ratio Test
Asymptotic Normality of Z^*



Ratio Test
Asymptotic Normality of Z^*



Ratio Test

Asymptotic Normality of Z^*

- Interim statistic:

$$Z_k^* = \frac{\log(\hat{\delta} + k)^2 - \log(\delta_0 + k)^2}{\text{s.e.}(\log(\hat{\delta} + k)^2)}$$

- Z_k^* is approximately normally distributed, and

$$Z_k^* \to Z^* \quad (\text{when } k \to \infty)$$

Ratio Test

Asymptotic Normality of Z^*

- Z^* will quickly converge to the **standard normal distribution**, i.e.,

$$Z^* \sim N(0, 1)$$

Ratio Test

Asymptotic Normality of Z^*

Normality of Z^* (Xeloda trials, simulation runs=100,000)

Number of Events	600	800	1000	1200	1400	1600	1800
p	68.2%	80.9%	88.9%	93.8%	96.6%	98.2%	99.1%

where p = proportion of simulation runs passed Shapiro-Wilk test.

Two 95% CI Method

Two 95% CI method:

- Define the non-inferiority cutoff (1+margin) as

  1 + (0.5)(LL of 95% CI for HR(P/C) - 1).

- If the 95% CI for HR(T/C) lies entirely beneath this cutoff, non-inferiority is concluded.

Hasselblad & Kong

$$\hat{\delta} = \frac{\log \hat{HR}(P1/C1) - \log \hat{HR}(T/C2)}{\log \hat{HR}(P1/C1)}$$

A “95%” confidence interval is calculated using a normal distribution with standard error

$$\sqrt{\left(\frac{\log \hat{HR}(T/C2)}{\log \hat{HR}(P1/C1)}\right)^2 \left(\frac{Var(\log \hat{HR}(T/C2))}{\log \hat{HR}(T/C2)^2} + \frac{Var(\log \hat{HR}(P1/C1))}{\log \hat{HR}(P1/C1)^2}\right)}$$

Simon's Method

The posterior density for λ=logHR(T/C) is N~ with mean (μ+y) and variance ($\sigma^2+\tau^2$).

Y: log HR(T/C), μ: log HR(C/P)
σ^2: Var(log HR(T/C)) τ^2: VarlogHR(C/P)

The posterior prob (T is superior to C):

$$P(\lambda<0)=1-\Phi[(\mu+y)/\text{sqrt}(\sigma^2+\tau^2)]$$

The prob (1-k)100% of the effect of C to P is not lost with T is Pr($\lambda-k\beta<0$, $\beta<0$).

Example

Xeloda vs 5-FU+LV

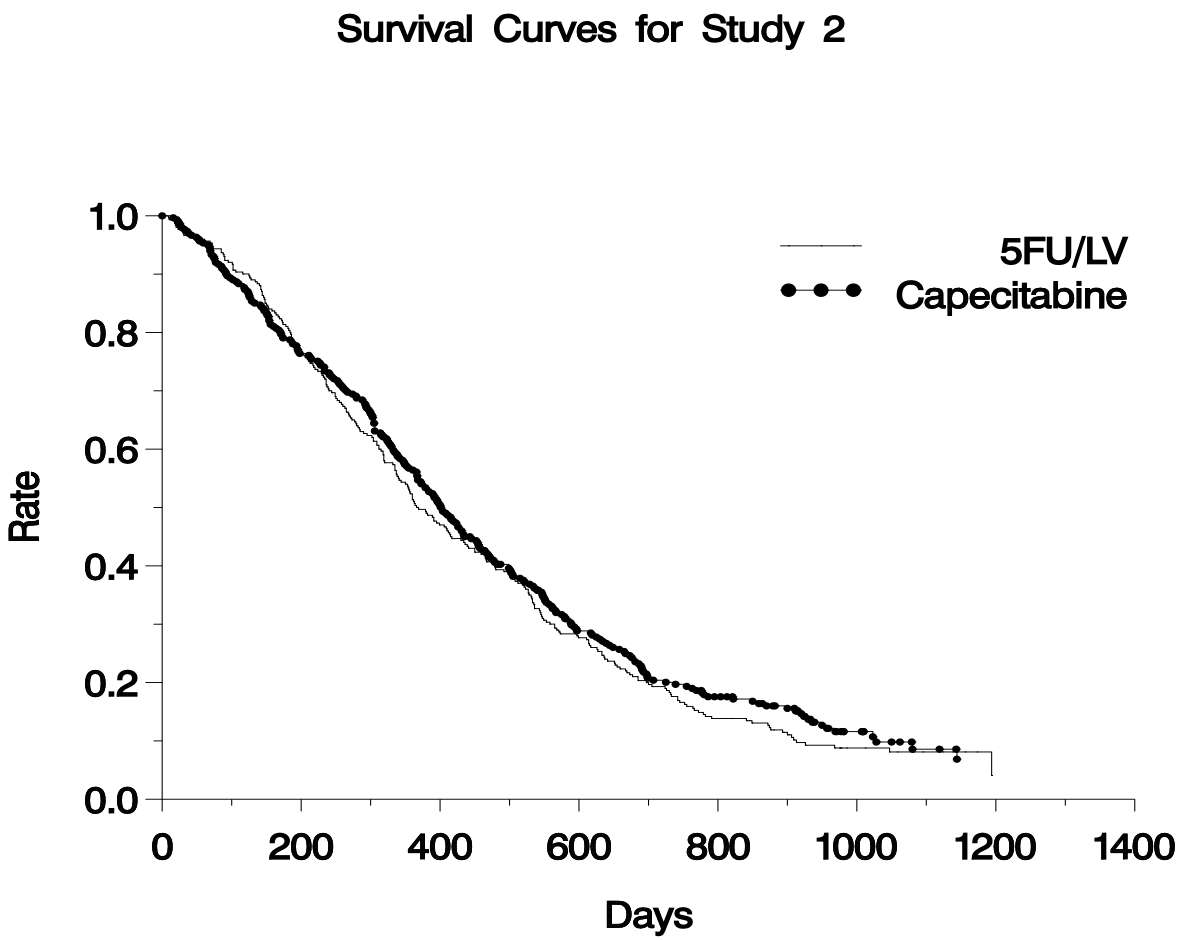
Xeloda trial

- Phase III Active Controlled Study
- Indication: First-line Metastatic Colorectal Cancer
- Rx: Xeloda (Capecitabine)
- Active Control: 5-FU+LV
- Primary endpoint: survival

Xeloda trial

	Xeloda	5FU/LV
Median Survival	13.5 (12.2-15.1)	12.3 (11.2-14.3)
Hazard Ratio: (Xeloda:5FU/LV)	0.92 (0.78-1.09)	

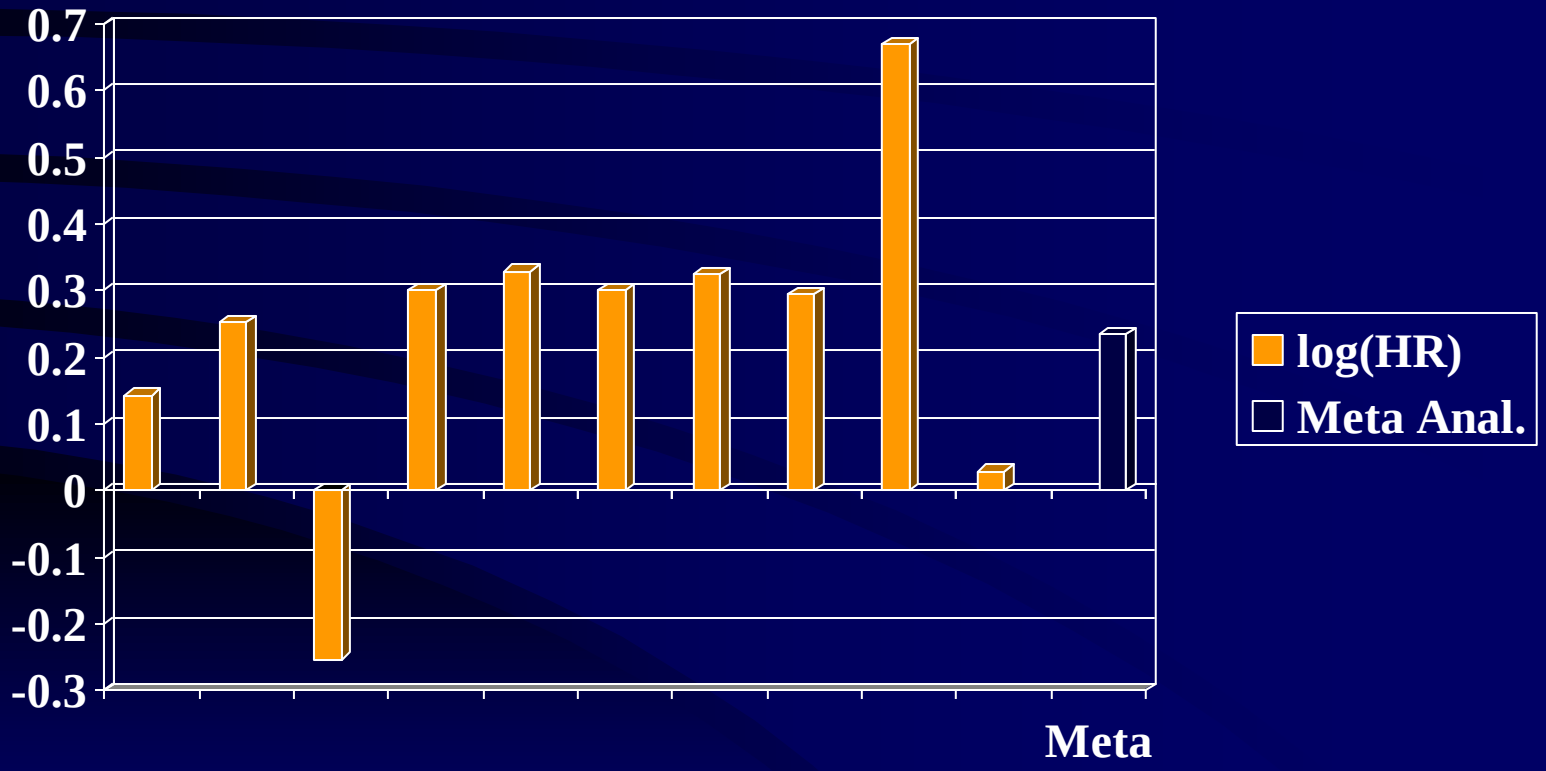
Xeloda trial



Active control effect

- Survival endpoint: HR(P/C)
- Multiple placebo controlled studies conducted for control effect
- Current trial population is similar to historical trial population(s)
- The effect size is not small.

Active control effect (5FU vs. 5FU/LV Trials)



Active control effect (5FU vs. 5FU/LV)

Random Effects Meta- analysis Model results based on ten trials

Summary of Relevant Survival Descriptive Statistics

HR(5-FU/5-FU+LV)	log HR	SE(logHR)
1.264	0.234	0.075

95% C.I. for HR(5-FU/5-FU+LV) is (1.091, 1.464)

Results of Xeloda and 5FU/LV trials

Xeloda trial:

- HR(T/C)=HR(Xeloda/5FU+LV)=0.92
- logHR(T/C)=-0.0844, SE(logHR)=0.087

Meta-analysis of 5FU/LV trials:

- HR(P/C)=1.264,
- logHR(P/C)=0.234, SE(logHR(P/C)=0.075

Linear Test

- δ defined using log HR, H_0: δ < 0.5, Z*=-2.13

Trial	δ^	p-value	Study Power	95% CI of δ^
Xeloda	**136.0%**	**0.0165**	**45.62%**	**(59.0%, 260%)**

Ratio Test

Trial	δ^	p-value	Study Power	95% CI of δ^
Xeloda	130.7%	0.0109	62.34%	(72.9%, 188%)

Two 95% CI Method

HR[1]	95% CI	Cutoff[2]	Fraction Demonstrated
0.92	0.78-1.09	1.046	2%

[1]HR: Hazard Ratio of Xeloda/5-FU/LV

[2]Cutoff for 50% retention.

Hasselblad & Kong's Method

- Estimated δ=1.36
- 95% CI is: 0.596-2.124

Bayesian Method
- Non-informative Priors

- Normal posterior probability distributions (or a posterior bivariate normal distribution) are determined from non-informative priors.
- A posterior probability is found for the event that both log HR(T/C2) < (1-δ)log HR(P1/C1) and log HR(P1/C1) >0. If this probability is greater than 0.975, non-inferiority is concluded.

Bayesian Method

- Joint Prob (logHR(T/C2)<(1-delta)logHR(P1/C1)) and logHR(C/P)>0 = 0.987.

Major issues

- The following are important design, conduct, analysis and interpretation issues
  - The choice of endpoints
  - The selection of the non-concurrent or historical studies
  - The modeling of the active control effect
  - The formulation of the hypotheses
  - The choice of fraction retention/margin
  - The interpretation of the results

Summary

- If control effect is small, active control trial should be a "superiority" trial, not a "non-inferiority" trial.
- Appropriate assessment of the control effect based on historical data may be difficult when
  - few trials
  - changing the population
  - changing the standard care
- Selection of the fraction retention should be based on both clinical and statistical judgment.
- Interpretation of results needs to be with caution.

END

Thanks