Design of Non-inferiority Trials – Problems, Current Approaches and Applications of Inferiority Index

George YH Chi, Ph.D.

Gang Li, Ph.D.

November 9, 2010

Statistical Science/Biostatistics & Programming
Global Development Organization
Johnson & Johnson PRD, L.L.C.

Content of Presentation

1. Definition of a NI trial
2. The Primary objective of a NI trial
3. Problems of and related to a NI trial
4. What should and can one do?
5. Two historical data based approaches to NI margin specification
6. Problems with the approaches
7. What is the best course forward?
8. Theory of inferiority index
9. The NI hypothesis and its test statistic
10. Time-to-event endpoint
11. Application to binary outcome
12. Summary, conclusion and recommendation
13. Theory of inferiority index – looking into the future

B&P BIOSTATISTICS & PROGRAMMING

1. Definition of a Non-Inferiority Trial

For this presentation, I will restrict non-inferiority trial (NIT) to mean a

- **Parallel randomized DB trial *without Placebo***
- **For ethical reason, if there is an available effective therapy, then that therapy should be used as the control instead of placebo (raison d'être)**

2. Primary Objective of a Non-Inferiority Trial

The primary objective of a NIT is to demonstrate the *effectiveness* of the new experimental treatment.

- **The regulation doesn't state how effectiveness is to be established, and hence does not explicitly require that the new experimental treatment be shown to be superior to a control**
- **FDA's past practices favors showing T > P.**
- **Thus, the implicit objective of a NIT is to show that T > "P", a virtual placebo that is not present.**

3. Problems of and Related to NIT

There are three different types of problems of and related to a NIT :

- **Inherent and irresolvable problems in a NIT**
- **Problem related to margin specification**
- **Impact on future products – the problem of bio-creep (efficacy and/or safety), a regulatory concern**

3. Problems of and Related to NIT (cont.)

Inherent and irresolvable problems of a NIT :

Absence of a placebo ⇒

- **Unable to determine the effectiveness of the control (relative to placebo) in the current trial**
- **Unable to determine if the current trial has sufficient assay sensitivity to be able to detect an expected control effect if present**

3. Problems of and Related to NIT (cont.)

Problems related to margin specification in a NIT:

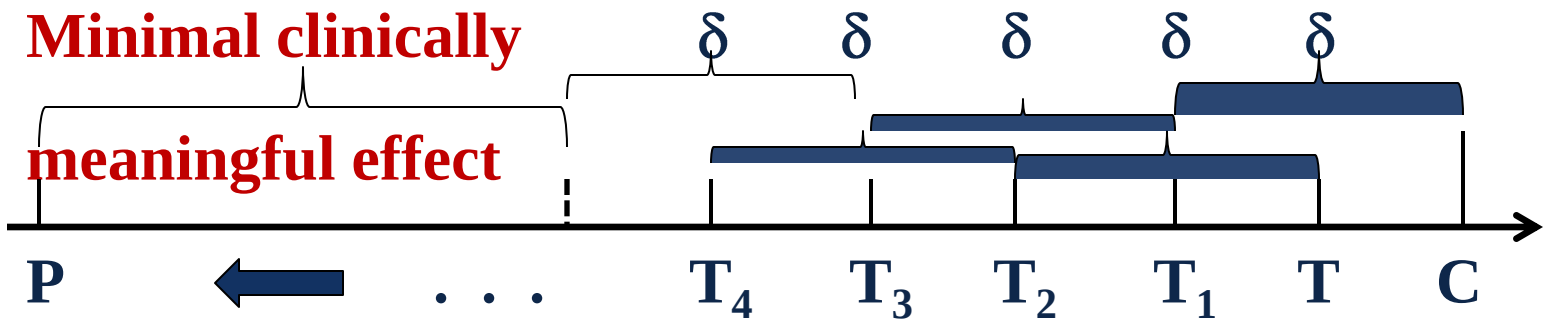
Assuming that the control has an effect in the current NIT and the NIT has the requisite assay sensitivity, then the two most often discussed approaches to the design of NIT have various problems:

- **Fixed margin approach**
- **Fraction retention (synthesis) approach**

3. Problems of and Related to NIT (cont.)

Impact on future products – Bio-creep:

If the new experimental drug T is shown to be non-inferior to the control C in the current NIT by clearing the specified margin δ as depicted in the diagram below, then the next product T1 can be compared to T, say using the same δ. The following diagram shows the potential loss of effect:



4. What should and can one do?

1. **First ask questions about the necessity of doing a NIT:**

   - **Is it really unethical to use placebo?**
   - **If it is, then are there other appropriate alternative designs besides NIT?**
   - **If not, then what is the best available therapy that can be used as the control?**

4. What should and can one do? (cont.)

2. **If we have to do a NIT, then we must make the following two assumptions:**
   - **The constancy of the control effect**
   - **NIT has assay sensitivity**

But despite having to make the above assumptions, we also

- **Have to make an earnest effort to enhance and ensure the validity of these two assumptions in NIT**
- **From historical data (if available)**
  - **estimate historical control effect and expected control effect, which should exceed clinically meaningful effect size**
  - **Identify key features of study design that would ensure sufficient assay sensitivity and raise the likelihood of observing the expected control effect**

4. What should and can one do? (cont.)

3. **To determine an appropriate NI margin with minimal subjectivity based on:**
   - **Expected control effect from historical data**
   - **NIT has assay sensitivity to detect the expected control effect and hence also distinguish T from C if T is not effective.**
   - **"Consideration of bio-creep"**

   and

   - **Theory of inferiority index, which provides a standard for measuring the degree of tightness of a margin, as well as a method for specifying margin relative to a reference control without relying on historical data.**

5. Two historical data based approaches to NI margin specification

Let T and C denote experimental treatment and control respectively, X an endpoint, X_T and X_C outcomes on T and C respectively. Assume smaller value of X represents worse outcome. Consider $X_T \sim N(\mu_T, \sigma_T^2)$ and $X_C \sim N(\mu_C, \sigma_C^2)$.

1. **The Fixed Margin Approach**

$$H_o: \mu_T - \mu_C \leq -M \quad \text{vs.} \quad H_a: \mu_T - \mu_C > -M$$

where $M > 0$ is a fixed number to be determined as follows:

5. Two historical data based approaches to NI margin specification (cont.)

Steps required in determining the fixed margin:

Step 1: If there is no historical data, this approach stops here. Otherwise, identify a set of relevant and appropriate historical studies.

- **Perform an appropriate meta-analysis based on random effects model on the efficacy endpoint X.**
- **Obtain the estimate of $\mu_C - \mu_P$ and its 95% confidence interval. Let M_1 = lower limit of the 95% CI. Based on clinical judgment, determine if the control effect size M_1 is clinically meaningful. If not, then the control cannot be used. Otherwise, M_1 will be considered as the control effect expected to be observed in the current NIT. (The Constancy Assumption implies $M_1 = \hat{\mu}_C - \hat{\mu}_P$, so conservatism is applied here)**

Step 2: Based on clinical judgment and the effect size M_1, a fraction ϕ is chosen that represents the proportion of the M_1 that T is required to preserve. Let $M_2 = (1-\phi)M_1$. Then $M = M_2$ is the fixed margin.

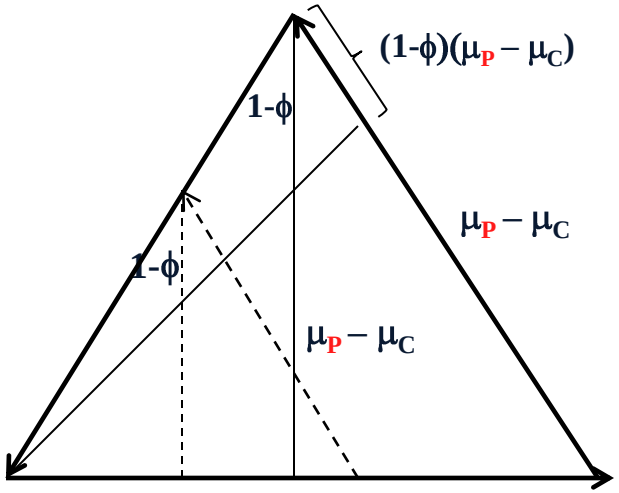
Step 3: Examine the historical studies and together with clinical knowledge, determine the key study features, aspects of study conduct, and data quality issues that need to be taken into account in the design and conduct of the current NIT in order to enhance the likelihood that the NIT has the required *assay sensitivity* so that the expected control effect M_1 determined above can be detected.

5. Two historical data based approaches to NI margin specification (cont.)

2. The Fraction Retention (Synthesis) Approach:

$H_o: (\mu_T - \mu_C)/(\mu_P - \mu_C) \geq 1-\phi$ vs. $H_a: (\mu_T - \mu_C)/(\mu_P - \mu_C) < 1-\phi$

where ϕ = the fraction of the control effect that the new experimental therapy T is expected to retain, and $(\mu_P - \mu_C) < 0$ as illustrated in Figure below:



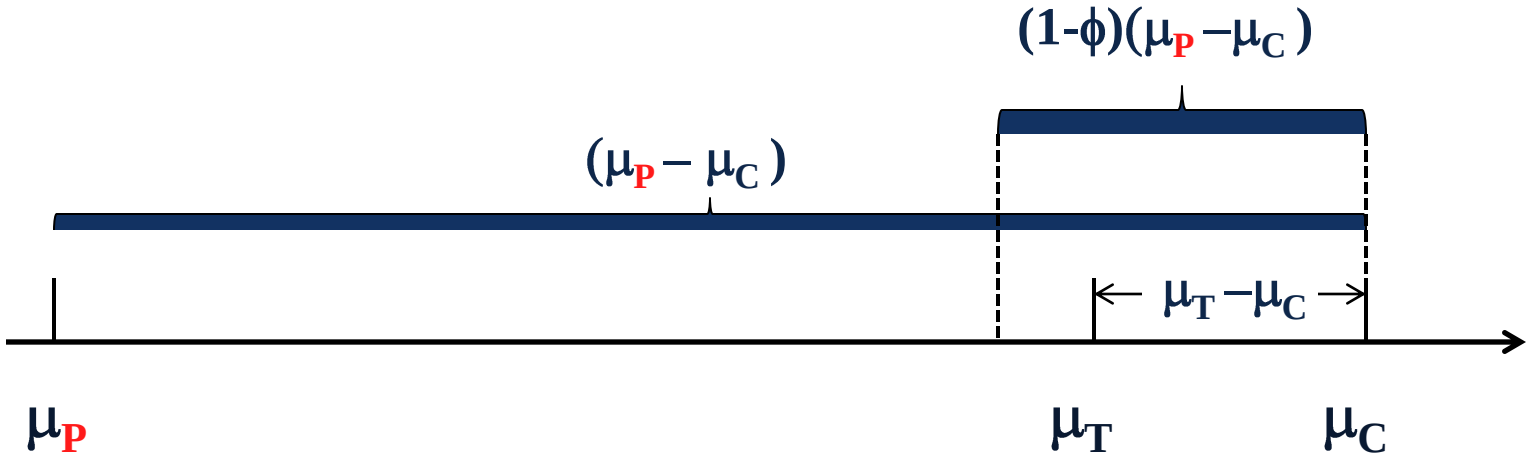
B&P BIOSTATISTICS & PROGRAMMING

5. Two historical data based approaches to NI margin specification (cont.)

2. **The Fraction Retention Hypothesis can be linearized as follows:**

$$H_o: (\mu_T - \mu_C) - (1-\phi)(\mu_P - \mu_C) \leq 0 \quad \text{vs.} \quad H_a: (\mu_T - \mu_C) - (1-\phi)(\mu_P - \mu_C) > 0$$

where ϕ = the fraction of the control effect that the new experimental therapy T is expected to retain and $\mu_P - \mu_C \neq 0$.



6. Problems with the two approaches

1. **Fixed Margin Approach: A basic *lack of objectivity* in the determination of NI margin M**

- **Inclusion/exclusion of historical data**
- **Choice of 95% confidence interval lower limit for the estimate of historical control effect (M_1)**
- **Fraction ϕ of M_1 to be retained ($M_2 = (1-\phi)\, M_1$)**

6. Problems with the two approaches (cont.)

2. **Fraction retention method**

- **Problem from the clinical perspective:**

  If the NI hypothesis assumes the true control effect to be unknown, then how can we assess how much of it to be retained, or allowed to be lost?

- **Problem from the statistical perspective**

  - Historical data come from studies that have been completed and have been repeatedly analyzed.
  - The synthetic test statistic does not provide valid statistical inference. The type I error rate is not controlled.

7. What is the best course forward?

Recommend fixed margin approach but minimize the subjectivity involved through an integrative and iterative application of *the theory of inferiority index* in conjunction with clinical knowledge and historical data

Inferiority Index is a very general concept

- **That can serve as a standard objective measure of the degree of stringency of a margin however it is derived.**
- **That can also specify a margin with any degree of stringency without relying on historical data**
- **That can provide valid statistical inference**

8. The Theory of Inferiority Index

The Theory of Inferiority Index

- **History**

  - **Equivalence margins or non-inferiority margins are usually defined in terms of differences between the means of the treatment T and the comparator C. The difference between T and C is measured by their mean difference.**
  - **There were early research publications that discuss differences or similarity between two T and C in terms of distributional differences:**

    - **The Dissimilarity Index:** $D = \frac{1}{2} \int_{-\infty}^{\infty} |f_T(t) - f_C(t)| \, dt$ **[Pearson (1895)]**
    - **The Overlapping Coefficient:** $OVL = \int_{-\infty}^{\infty} min\,[f_T(t), f_C(t)] \, dt$ **[Weitzman (1970)]**
    - **Separation Coefficient** $Cs = 1 - 2\Phi(-(|\mu_T - \mu_C|/2\sigma)$ **[Jensen (1991)]**

8. The Theory of Inferiority Index

The Theory of Inferiority Index

- **History (Continuation)**

  - **Even though Jensen indicated that in general the overlapping coefficient OVL appears to provide more insights into the degree of difference that exists between two distributions than the dissimilarity index D, we believe that OVL will not provide the kind of measure of *degree of inferiority* that we need in margin specification.**

  - **In anthropometry, Mora (1989) proposed a measure of nutritional deficiency between a study population and a reference population in terms of the total area under the study population density function that is outside the reference population density function.**

  - **Bohning et al. modified Mora's definition and considered the non-parametric statistical inference based on this modified measure. They further noted the link between their measure and the Kolmogorov-Smirnov statistic.**

8. The Theory of Inferiority Index

The Theory of Inferiority Index

- **Definition of Inferiority Index**
  - **The inferiority index of T relative to C is defined as**

$$\rho = \sup_{-\infty<t<\infty} [F_T(t) - F_C(t)]$$

For normal distribution, let $X_T \sim N(\mu_T, \sigma_T^2)$ and $X_C \sim N(\mu_C, \sigma_C^2)$, and assume that $\mu_T < \mu_C$.

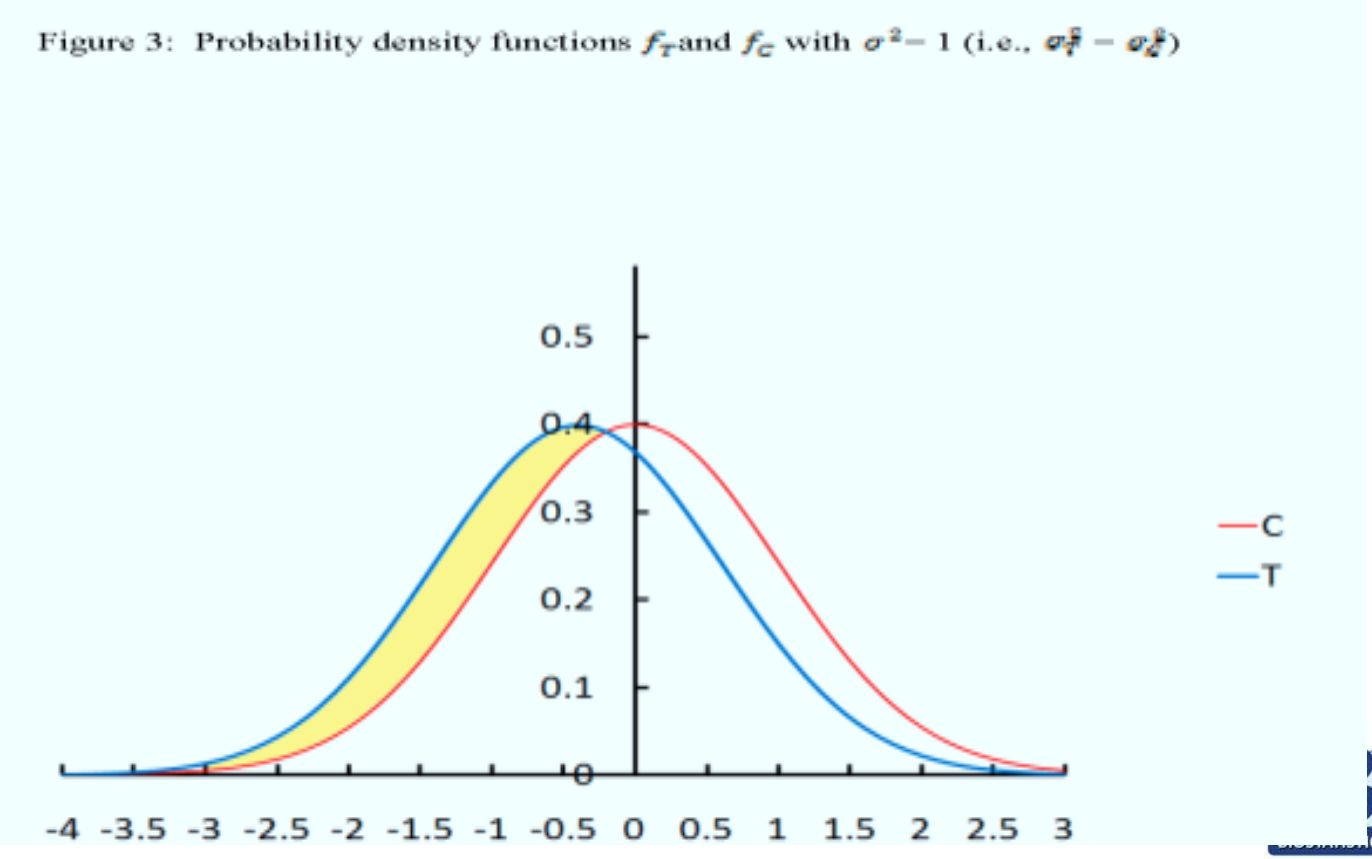
- **We first standardize both X_C and X_T relative to the scales of X_C and thus we have**

$$X_C \sim N(0,1) \text{ and } X_T \sim N(\mu, \sigma^2),$$

where $\mu = (\mu_T - \mu_C)/\sigma_C$ and $\sigma^2 = \sigma_T^2 / \sigma_C^2$

8. The Theory of Inferiority Index (cont.)

- The Theory of Inferiority Index
  - **Why It Measures Inferiority**
    - $\sigma_T^2 = \sigma_C^2$



8. The Theory of Inferiority Index (cont.)

- **The Theory of Inferiority Index**
  - **Why It Measures Inferiority (Continuation)**
    - $\sigma_T^2 < \sigma_C^2$

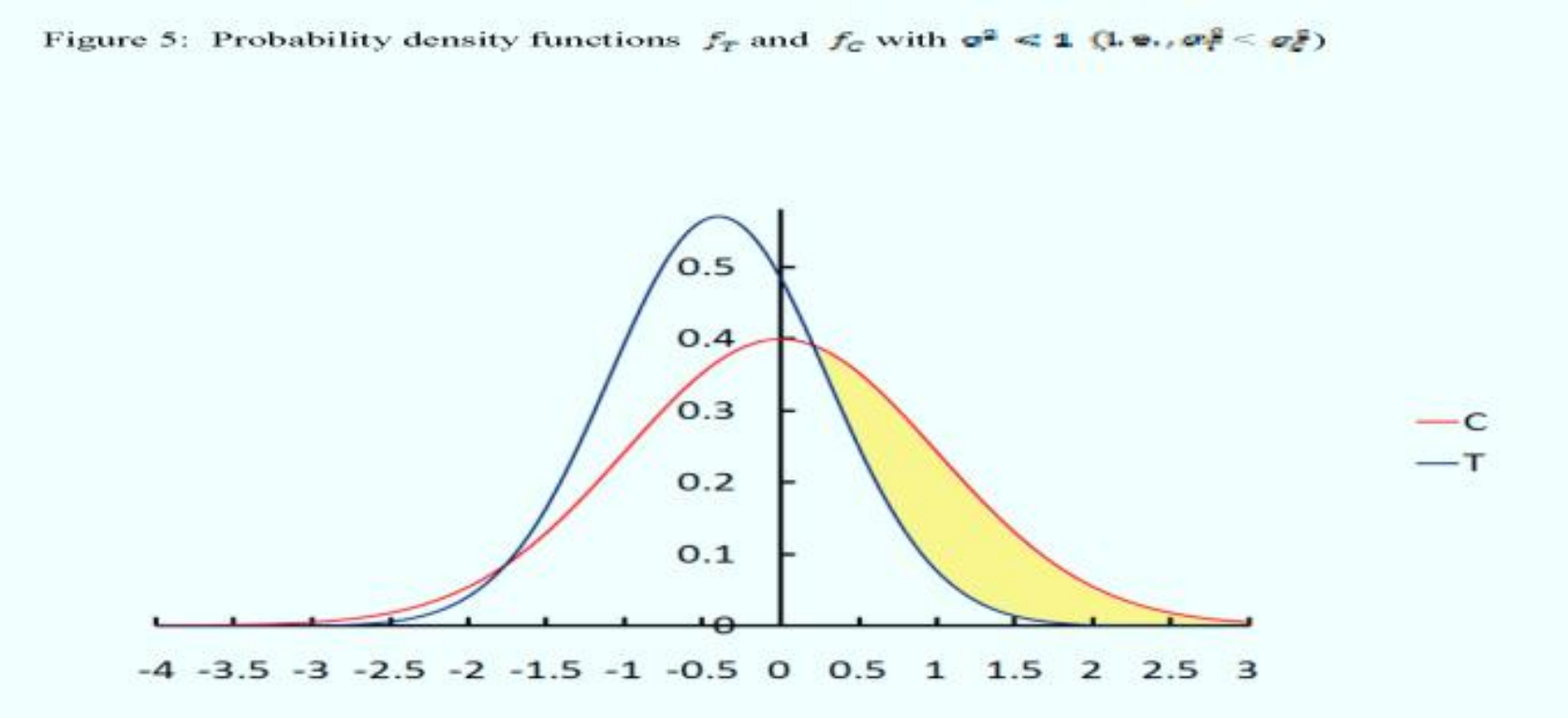
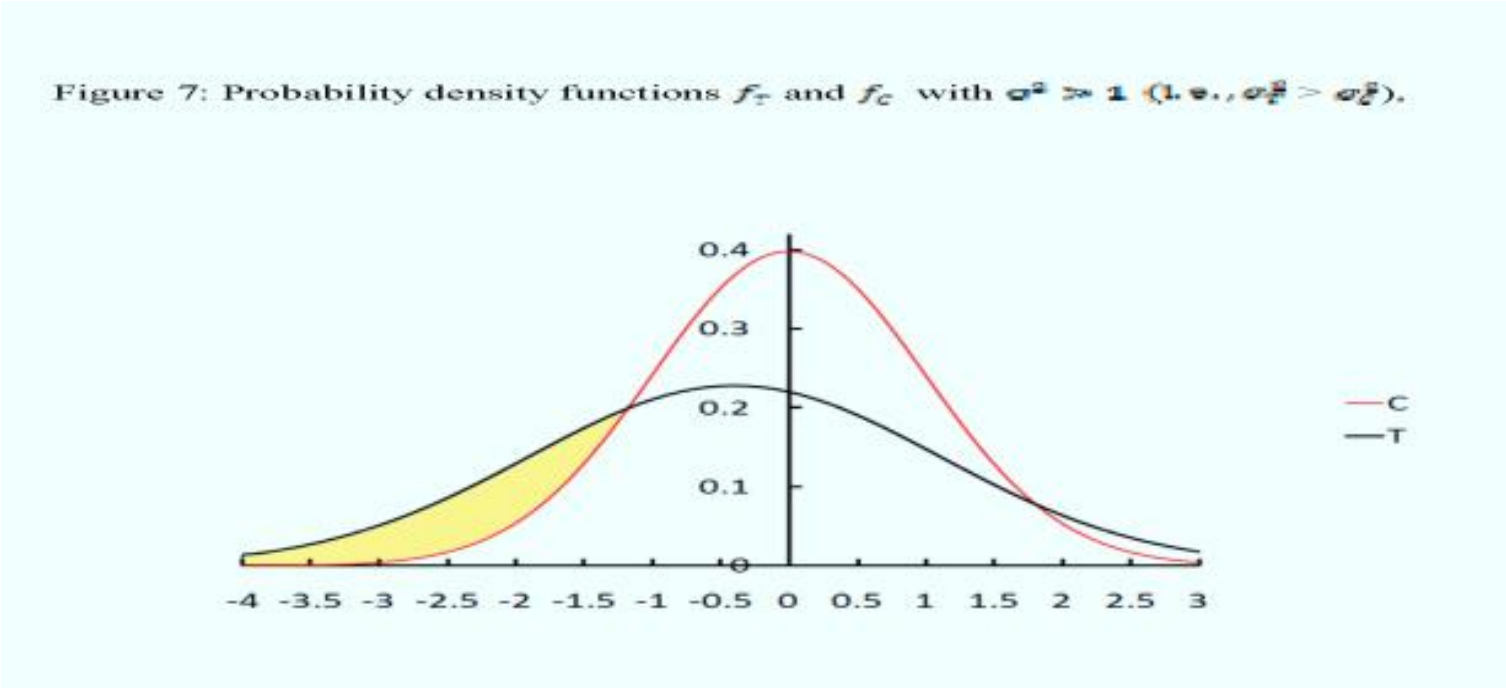


Figure 5: Probability density functions f_T and f_C with $\sigma^2 < 1$ (i.e., $\sigma_T^2 < \sigma_C^2$)

8. The Theory of Inferiority Index (cont.)

- **The Theory of Inferiority Index**
  - **Why It Measures Inferiority (Continuation)**
    - $\sigma_T^2 > \sigma_C^2$



B&P BIOSTATISTICS & PROGRAMMING

8. The Theory of Inferiority Index (cont.)

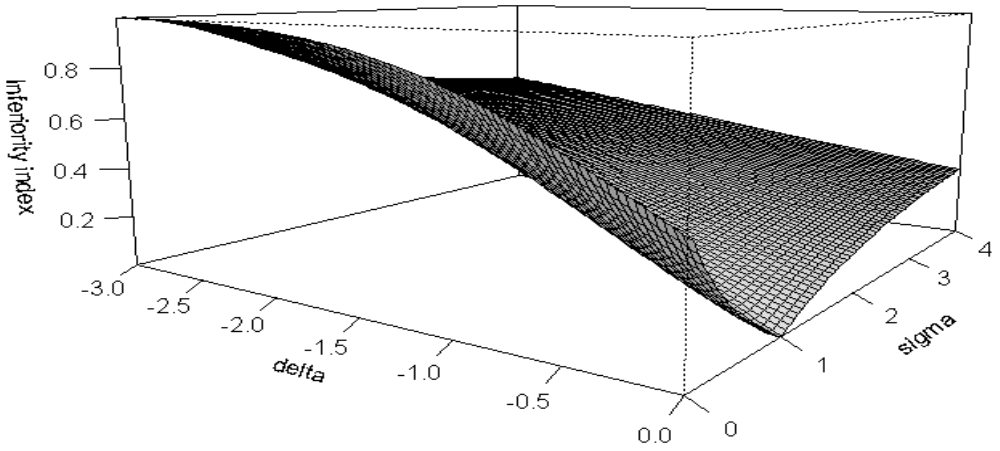
The Theory of Inferiority Index

Specification of the Non-inferiority Margin

- For given $\delta < 0$ and σ^2 there corresponds a unique inferiority index $\rho = g(\delta, \sigma)$
- Conversely, for a given ρ and variance ratio σ^2 in certain restricted interval, there corresponds a unique $\delta(\rho,\sigma) = g^{-1}(\rho,\sigma) < 0$.
- Since $\delta = (\mu_T - \mu_C)/\sigma_C$, $\delta(\rho,\sigma)$ can be used to specify the non-inferiority margin with a desired degree of inferiority index ρ with a given variance ratio σ^2.

8. The Theory of Inferiority Index (cont.)

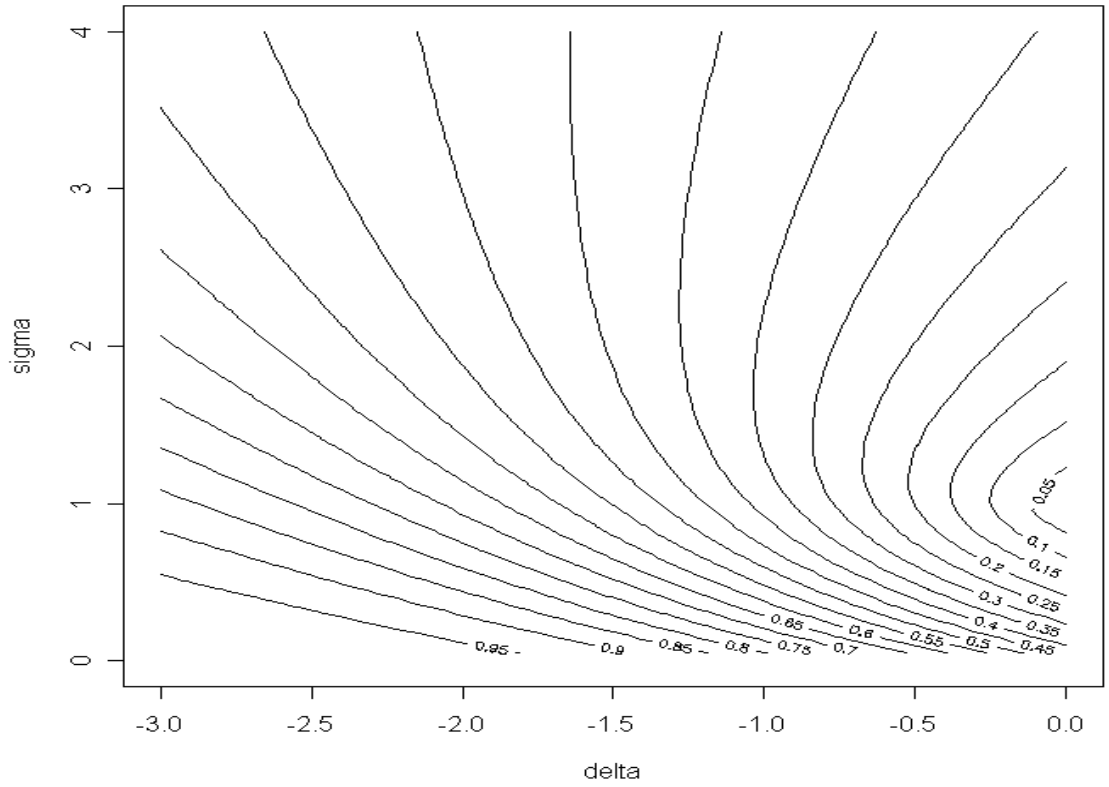
- **The Theory of Inferiority Index**
  - **Specification of the Non-inferiority Margin**



B&P BIOSTATISTICS & PROGRAMMING

8. The Theory of Inferiority Index (cont.)

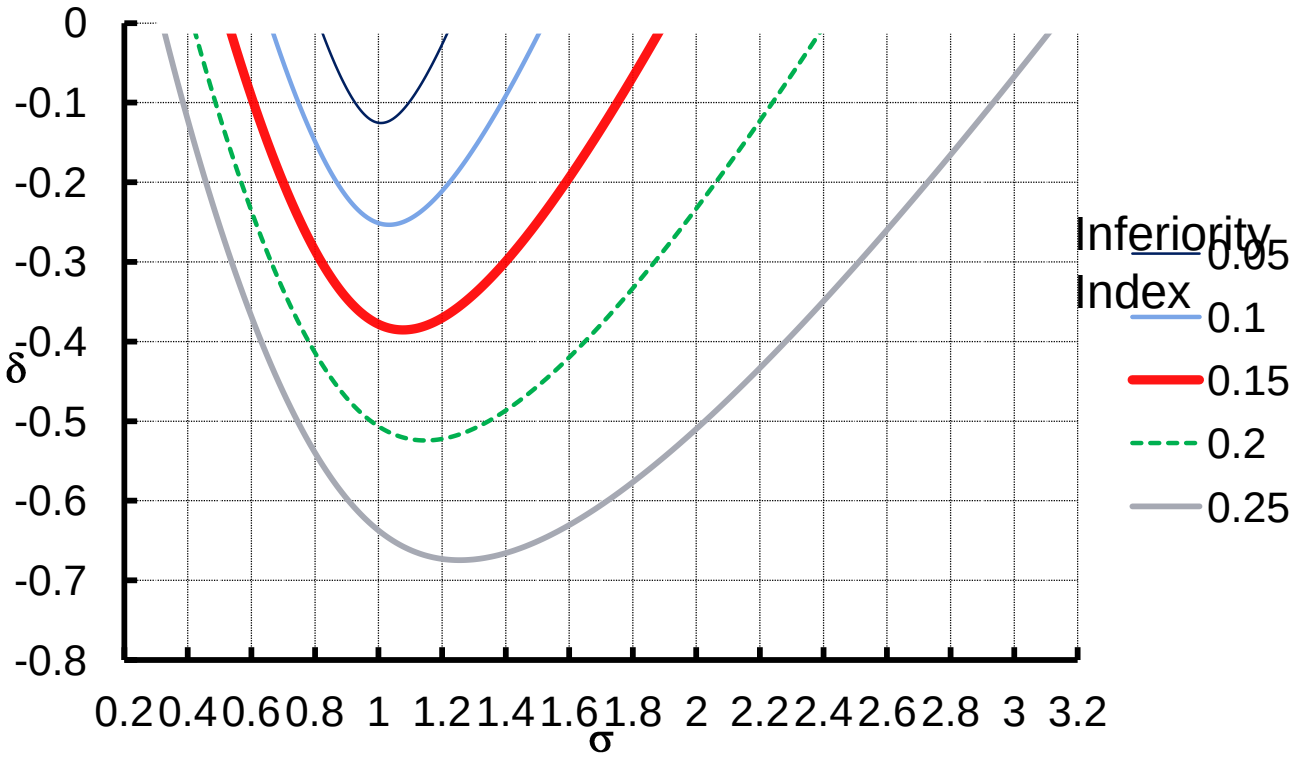
- **The Theory of Inferiority Index**
  - **Specification of the Non-inferiority Margin**



B&P BIOSTATISTICS & PROGRAMMING

8. The Theory of Inferiority Index (cont.)

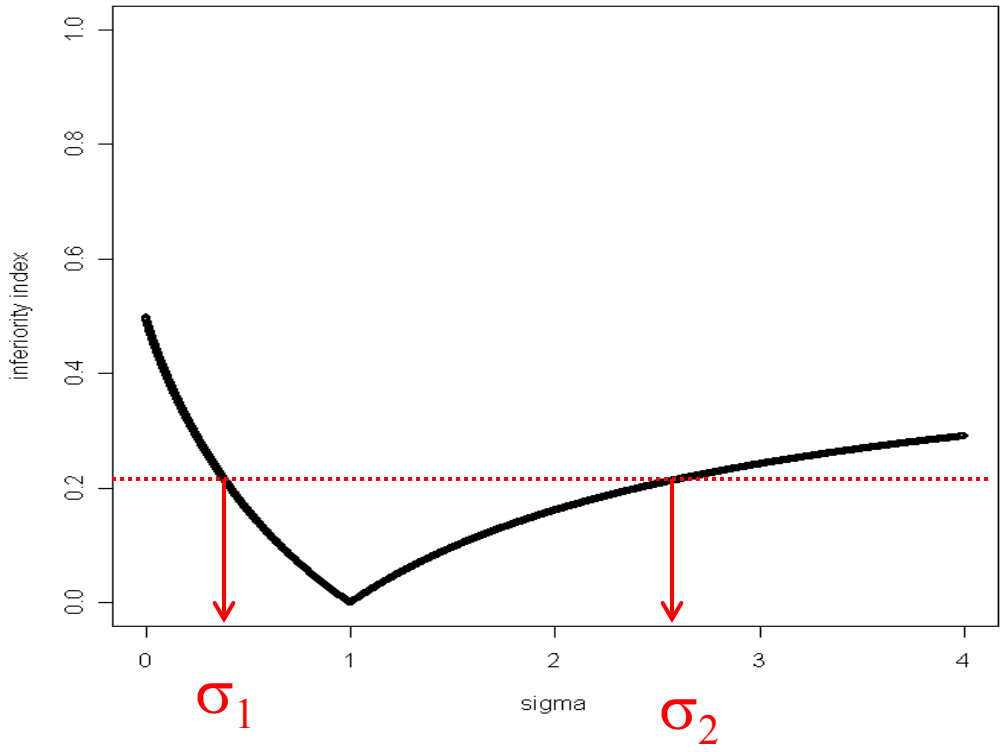
- **The Theory of Inferiority Index**
  - **Specification of the Non-inferiority Margin**



B&P BIOSTATISTICS & PROGRAMMING

8. The Theory of Inferiority Index (cont.)

- **The Theory of Inferiority Index**
  - **Specification of the Non-inferiority Margin**



9. The NI Hypothesis and its Test Statistic

The Non-inferiority Hypothesis

- **Therefore, in designing a NI trial, for a given inferiority index ρ and a given variance ratio σ^2 in certain restricted interval, there exists an non-inferiority margin $\delta(\rho,\sigma) = g^{-1}(\rho,\sigma) < 0$. We can define the non-inferiority hypothesis as:**

$$\text{Ho: } (\mu_T - \mu_C)/\sigma_C \leq \delta(\rho,\sigma) \quad \text{vs.} \quad \text{Ha: } (\mu_T - \mu_C)/\sigma_C > \delta(\rho,\sigma)$$

9. The NI Hypothesis and its Test Statistic (cont.)

The Test Statistic

- **A General Theorem**

Let $X_T \sim N(\mu_T, \sigma_T^2)$ and $X_C \sim N(\mu_C, \sigma_C^2)$, and assume that $\mu_T < \mu_C$. Let ρ be a given inferiority index and σ^2 a given variance ratio. Then, there corresponds a unique $\delta(\rho,\sigma) = g^{-1}(\rho,\sigma) < 0$.

Let the NI hypothesis be defined by:

$$\text{Ho: } (\mu_T - \mu_C)/\sigma_C \leq \delta(\rho,\sigma) \quad \text{vs.} \quad \text{Ha: } (\mu_T - \mu_C)/\sigma_C > \delta(\rho,\sigma)$$

A test statistic for the above NI hypothesis is given by:

$$\hat{\delta} = \sqrt{n}\left[\frac{(\hat{X}_T - \hat{X}_C)}{\hat{\sigma}_C} - \delta(\rho,\sigma)\right] \rightarrow N(0,\ S),\ \text{where } \hat{\sigma}_C$$

$$\text{is an estimate of } \sigma_C, \text{ and } S = \{(1+\sigma^2) + \frac{\delta^2(\rho,\sigma)}{4}\}.$$

$$\text{The NI hypothesis is rejected if } \frac{\sqrt{n}\left[\frac{(\hat{X}_T - \hat{X}_C)}{\hat{\sigma}_C} - \delta(\rho,\sigma)\right]}{\sqrt{S}} > z_{\alpha/2}.$$

9. The NI Hypothesis and its Test Statistic (cont.)

Some Special Cases

- **Binary Outcome**

Let $X_T \sim \text{Bernoulli}(\pi_T, \sigma_T^2)$ and $X_C \sim \text{Bernoulli}(\pi_C, \sigma_C^2)$, and assume that $\pi_T < \pi_C$. Let ρ be a given inferiority index and σ^2 a given variance ratio. Then, there corresponds a unique $\delta(\rho,\sigma) = g^{-1}(\rho,\sigma) < 0$.

Let the NI hypothesis be defined by:

$$\text{Ho: } (\pi_T - \pi_C)/\sigma_C \leq \delta(\rho,\sigma) \quad \text{vs.} \quad \text{Ha: } (\pi_T - \pi_C)/\sigma_C > \delta(\rho,\sigma)$$

A test statistic for the above NI hypothesis is given by:

$$\hat{\delta} = \sqrt{n}\left[\frac{(\hat{X}_T - \hat{X}_C)}{\hat{\sigma}_C} - \delta(\rho,\sigma)\right] \rightarrow N(0, S),$$

where $S = (1+\sigma^2)\{1 + \frac{\delta^2(\rho,\sigma)}{16\sigma_C^2\sigma^2}\}$, is estimated by

$$\hat{S} = (1+\sigma^2)\{1 + \frac{\delta^2(\rho,\sigma)}{16\hat{\sigma}_C\sigma^2}\}.$$

10. Definition of Inferiority Index – Time-to-Event

Survival Distribution:

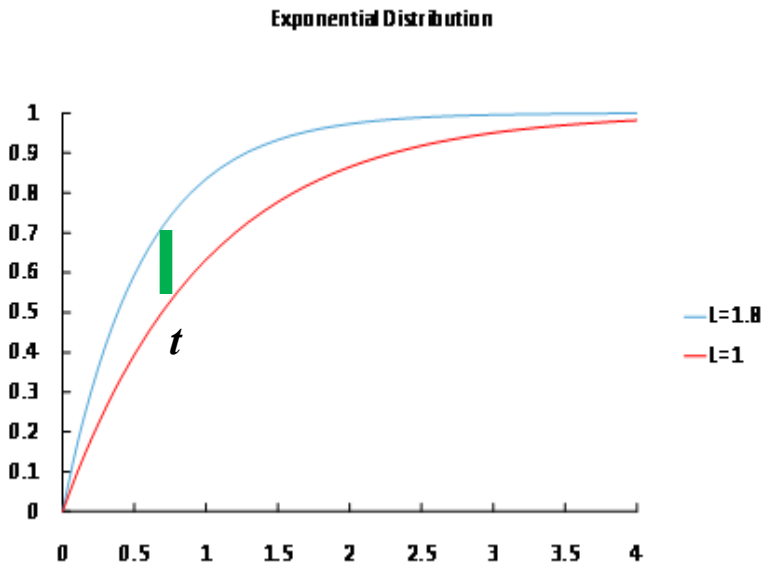
- Consider a trial to evaluate a treatment *T*, control *C*
- A primary endpoint *X* such that a smaller value represents a worse response

- The cumulative distribution functions are $F_T(t)$ and $F_C(t)$, respectively

- $F_T(t) - F_C(t)$ measures the probability that the treatment *T* is worse than the control at *t*, is an indicator of T inferior to C upto

- The *inferiority index* is defined as

$$\rho = Sup_{-\infty<t<\infty}[F_T(t) - F_C(t)]$$



10. Time-to-Event Endpoint (cont.)

- **Survival Trials**

For most survival trials, the proportional hazards assumption is made.

Under this assumption, our theory is much simpler and more elegant:

Let $X_T \sim S_T(t)$ and $X_C \sim S_C(t)$ be their corresponding survival distributions with their respective hazard rate h_T and h_C.

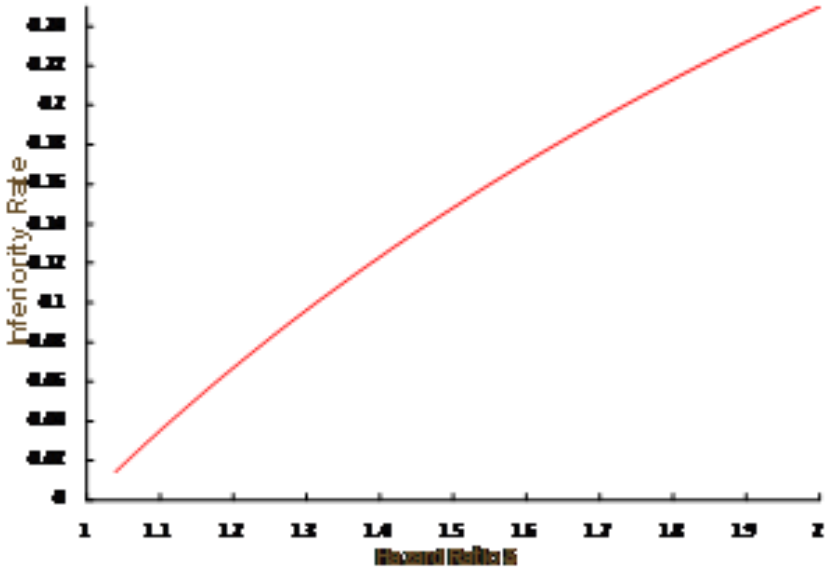
Under the Cox proportional hazard assumption, let $\delta = h_T/h_C$ be the Constant hazard ratio. Then, for each constant hazard $\delta = h_T/h_C$, there corresponds a unique inferiority index ρ, and conversely, for each given inferiority index ρ, there is a unique δ.

10. Time-to-Event Endpoint (cont.)

Theorem 1: Under the proportional hazards assumption, for each $\delta = \frac{h_T(t)}{h_C(t)} \geq$ corresponds a unique inferiority index

$$\rho = g(\delta) = e^{-\frac{\log \delta}{\delta - 1}}\left(1 - \frac{1}{\delta}\right), \text{ for } 1 \leq \delta < \infty;$$

g is a strictly increasing and continuously differentiable. Conversely, given an inferiority index, ρ, there is a unique $\delta = g^{-1}(\rho)$



10. Time-to-Event Endpoint (cont.)

- **Survival Trials (Continuation)**

Table: Selected Values of the Inferiority Index and the Corresponding Margin

ρ	0.018	0.035	0.050	0.067	0.096	0.213
δ	1.05	1.10	1.15	1.20	1.30	1.80

10. Time-to-Event Endpoint (cont.)

- **Survival Trials (Continuation)**

Example 1: In the FDA Draft Non-inferiority Guidance, for the example on TPA trial, a non-inferiority margin of 1.14 was used for the hazard ratio, while the sponsor proposed a non-inferiority margin of 1.16. Table above suggest that these margins correspond approximately to an inferiority index of ρ = 0.05 based on a margin of $g^{-1}(0.05) = 1.15$.

Example 2: In the FDA Draft Non-inferiority Guidance, for the fourth example on Xeloda, the FDA reviewer applied meta-analysis on historical studies to derive a non-inferiority margin of 1.09, which is the lower limit of the 95% confidence interval for the control effect based on a mixed effects model analysis of the survival data from 10 historical studies comparing 5-FU+LV vs. 5 FU. According to the Table above, this margin corresponds to a inferiority index of ρ = 0.035.

10. Time-to-Event Endpoint - Example 1

- **Xeloda to Treat Metastatic Colorectal Cancer**

Study	HR (Xeloda/5-FU+LV)	95% CI
1	1.00	(0.84, 1.18)
2	0.92	(0.78, 1.09)

  - Based on NI synthesis test with 50% retention (can be translated to an approximate fixed margin of 1.107 for M_2): Study 1 failed; Study 2 met NI.
  - Based on a fixed margin approach M_1=1.09, under 50% retention rate M_2=1.045, so none of the studies met NI criterion
  - The NI margins for both approaches have very low inferiority index

δ	1.107	1.045
ρ	0.038	0.016

10. Time-to-Event Endpoint (cont.)

- **Survival Trials (Continuation)**

Example 3: In the recent December 2008 FDA Guidance for Industry on Diabetes Mellitus – Evaluating Cardiovascular Risk in New Anti-diabetic Therapies to Treat Type 2 Diabetes, the drug sponsors are asked to provide at the time of NDA submission evidence of cardiovascular safety by demonstrating that the hazard ratio of the new therapy relative to placebo for the composite cardiovascular endpoint of mortality, myocardial infarction, stroke, and may include hospitalization for acute coronary syndrome (ACS) and possibly other relevant clinical endpoints excludes a margin of 1.8 and then later a margin of 1.3. These safety margins correspond to inferiority index of

$$\rho = g(1.8) = 0.22 \text{ and } \rho = g(1.3) = 0.096$$

10. Time-to-Event Endpoint - Example 2 (cont.)

- FDA's guidance on 'Diabetes Mellitus – Evaluating Cardiovascular Risk in New Anti-diabetic Therapies to Treat Type 2 Diabetes' requires the CV risk ratio of the new treatment relative to control not too high:
  - To rule out risk ratio 1.8 at submission
  - To rule out risk ratio 1.3 at final

- The corresponding **ρ** of these margins are

	Submission	Final
δ	1.8	1.3
ρ	0.213	0.096

11. Application to Binary Outcome

- **Binary Outcome Trials**

Example 4: The specification of margin for binary outcomes has often been a problem, particularly in the area of anti-infective trials, because there is a lack of relevant historical studies. In addition, for binomial probability, if the new treatment is to show non-inferiority, it has to assume that its response probability is worse than the control, and hence for sure the variance of T and C will be different.

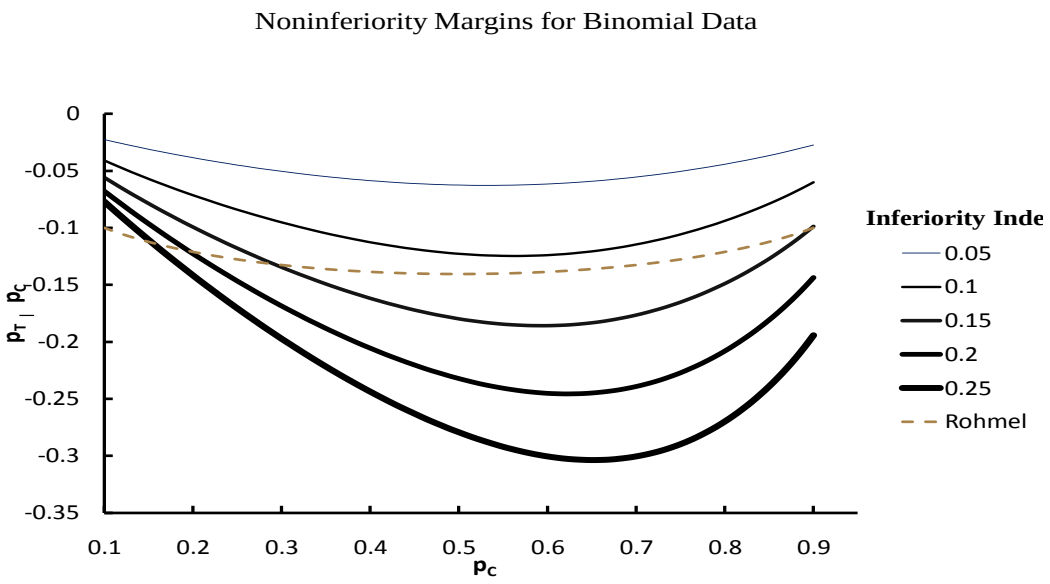
The 1992 FDA Anti-infective Points to Consider Guidance defines three margins for relative difference as follows: $\delta = -0.20$ for $\pi_C < 0.80$, $\delta = -0.15$ for $0.80 \leq \pi_C < 0.90$, and $\delta = -0.10$ for $\pi_C > 0.90$.

John Lewis had suggested that the margin should be a continuous function of the control response rate. Röhmel (2001) had come up with a continuous function defined as $\delta = g(\pi_C) = \pi_C - 0.223\{\pi_C(1-\pi_C)\}^{1/3}$. However, the rationale provided is based on a regulatory perspective and not statistically justified.

11. Application to Binary Outcome (cont.)

- **Binary Outcome Trials (Continuation)**

Based on our methodology, we are able to derive naturally a continuous curve for the margin for each level of inferiority index specified as shown in the following Figure.:



11. Application to Binary Outcome (cont.)

- **Binary Outcome Trials (Continuation)**

The following table presents NI margin values for relative difference $\Delta = P_T - P_C$ corresponding to a few selected values of inferiority index ρ and control rate P_C

Table 1: Relative Difference Margin Δ for Given ρ and P_C

	Inferiority Index (ρ)			
P_C	0.075	0.10	0.15	0.20
0.99	- 0.0066	- 0.0096	- 0.0172	- 0.0276
0.98	- 0.0120	- 0.0173	- 0.0303	- 0.0473
0.95	- 0.0255	- 0.0360	- 0.0610	- 0.0916
0.90	- 0.0430	-0.0601	-0.0987	-0.1435
0.85	- 0.0570	-0.0789	-0.1271	-0.1808
0.80	- 0.0684	-0.0940	-0.1489	-0.2080
0.75	- 0.0775	-0.1058	-0.1652	-0.2271

11. Table of Relative Difference Margin for Selected ρ

Table of Relative Difference Margins for Selected Inferiority Indices and Control Response Rates

Control Response Rate (Pc)	Inferiority Index 0.050	0.075	0.100	0.125	0.150	0.175
0.300	-0.05034	-0.07347	-0.09520	-0.11554	-0.13449	-0.15210
0.350	-0.05489	-0.08049	-0.10476	-0.12767	-0.14918	-0.16931
0.400	-0.05846	-0.08615	-0.01127	-0.13787	-0.16178	-0.18430
0.450	-0.06097	-0.09034	-0.01187	-0.14601	-0.17207	-0.19683
0.500	-0.06234	-0.09291	-0.12279	-0.15179	-0.17976	-0.20657
0.550	-0.06249	-0.09372	-0.12461	-0.15493	-0.18448	-0.21308
0.600	-0.06138	-0.09264	-0.12393	-0.15507	-0.18578	-0.21584
0.650	-0.05899	-0.08960	-0.12067	-0.15197	-0.18325	-0.21428
0.700	-0.05532	-0.08456	-0.11464	-0.14536	-0.17649	-0.20781
0.750	-0.05038	-0.07750	-0.10577	-0.13505	-0.16515	-0.19590
0.800	-0.04415	-0.06837	-0.09395	-0.12083	-0.14888	-0.17799
0.850	-0.03655	-0.05700	-0.07891	-0.10231	-0.12711	-0.15330
0.900	-0.02735	-0.04302	-0.06010	-0.07866	-0.09870	-0.12032

11. Application to Binary Outcome (cont.)

Assessing Low Incidence Rates

Example 1: Suppose that relative to certain known risk, a control is assumed to have an incidence rate of 2%, and based on clinical judgment, a new treatment should not incur an incidence rate in excess of more than 3%. That is,
$\delta = \Delta/\sigma_C$ = 0.03/sqrt(0.98*(1-0.98)) = 0.03/0.14 = 0.2143.

Note that in the formulation of our NI hypothesis, we assume that $P_T < P_C$, i.e., P_T is inferior to P_C. Hence, we should consider $P_T = 0.969$ and $P_C = 0.979$.

Here are the outcome of the trial data relative to this risk:

	T	C
Yes	8 (3.1%)	5 (2.1%)
No	252 (96.9%)	235 (97.9%)
Total	260	240

$T = (P_T-P_C)/\sigma_C\hat{} + \delta)/sqrt(2\hat{S}/(n_C+n_T)) = 1.16 \not> 1.96$
$\delta\hat{} = (P_T-P_C)/\sigma_C\hat{} - 1.96\ sqrt(2\hat{S}/(n_C+n_T)) = -0.3082 \not> -0.2143$
Hence, non-inferiority is not concluded for the given margin of 3%
$\delta\hat{}$ is the minimum δ that the data can clear NI, and it corresponds to an inferiority index of $\rho\hat{} \approx 0.515$.

11. Application to Binary Outcome (cont.)

Assessing Cure Rates

Example 2: Suppose we are designing a NI trial to compare the cure rate of a new anti-infective treatment T to an active control C. There is no historical study Comparing the control C to placebo. It is thought that the cure rate for C is about 80%. According to the current guideline, a NI margin of $\delta = -0.15$ or $\delta = -0.20$ should be used, since 0.80 sits right at the boundary between the two categories of cure rates: $P_C < 0.80$ and $0.80 \le P_C < 0.90$. Should we use – 0.20, or – 0.15?

Similarly, what if the cure rate for C is $P_C = 0.85$, 0.95? From our theory of inferiority index, we have the following $\Delta = P_T - P_C$ margins and inferiority indices:

Table 2

	Cure Rate of P_C			
ρ	0.75	0.80	0.85	0.95
0.10	− 0.1058	− 0.0940	− 0.0789	− 0.0360
0.15	− 0.1652	− 0.1489	− 0.1271	− 0.0610
0.20	− 0.2271	− 0.2080	− 0.1808	− 0.0916

12. Summary, Conclusion and Recommendation

- The theory of inferiority index is a very general concept for comparing two distributions
- It provides a standard objective measure of the degree of stringency of any margin however it is derived.
- It provides an objective way of margin specification without dependence on historical trials
- It allows a consistent, transparent, objective way to margin specification
- It should facilitate the regulatory agencies in the determination of specific margins for individual therapeutic areas in their guidance documents

13. Theory of Inferiority Index – Looking Into the Future

In view of its generality, inferiority index can be applied to any comparative study involving direct comparison

Examples:

- **Non-inferiority trials**
- **Bioequivalence studies**
- **Biosimilarity studies**
- **Safety assessment studies**
- **Comparative effectiveness studies**
- **Therapeutic equivalence (e.g., equi-analgesic trials)**
- **Anthropometric studies**
- **Defining responder relative to a reference population**

❖ References

- Böhning, D, Hempfling, A, Schelp, F-P, and Schlattmann, P (1992). The area between curves (ABC) – Measures in nutritional anthropometry, Statistics in Medicine 11: 1289-1304.
- Bradley, EL (1985). "Overlapping Coefficient", Encyclopedia of Statistical Sciences.
- Brown, D, Volkers, P, Day, S (2006). An introductory note to CHMP guidelines: choice of the non-inferiority margin and data monitoring committees, Statistics in Medicine 25: 1623-1627.
- Chi, G, Chen, G, Rothmann, M and Li, N (2010). Active Control Trials, Encyclopedia of Biopharmaceutical Statistics, Marcel Dekker, Inc. 9-15.
- Committee for Medicinal Products for Human Use (CHMP) (2005). Guideline on the choice of the non-inferiority margin, European Medicine Agency, European Medicines Agency, July 27, 2005.
- D'Agostino, RB (2003). Editorial - Non-inferiority trials: advances in concepts and methodology, Statistics in Medicine 22: 165-167.
- D'Agostino, RB, Sr., Massaro, JM, Sullivan, LM (2003). Non-inferiority trials: design concepts and issues – the encounters of academic consultants in statistics, Statistics in Medicine 22: 169-186.
- Dennis, NR and Carter, CO (1978). Use of overlapping normal distributions in genetic counseling, Journal of Medical Genetics 15: 106-108.
- Garbe, E, Röhmel, J and Gundert-Remy, U (1993). Clinical and statistical issues in therapeutic equivalence trials, European Journal of Clinical Pharmacology 45: 1-7.
- Garrett, AD (2003). Therapeutic equivalence: fallacies and falsification, Statistics in Medicine 22: 741-762.
- Hauck, WW, Chen, ML, Hyslop, T, Patnaik, R, Schuirmann, D, Williams, R (1996). Mean difference vs. variability reduction: tradeoffs in aggregate measures for individual bioequivalence, International Journal of Clinical Pharmacology and Therapeutics 34 912): 535-541.

B&P BIOSTATISTICS & PROGRAMMING

❖ References

- Hauck, WW, Parekh, A, Lesko, LJ, Chen, ML and Williams, RL (2001). Limits of 80% - 125% for AUC and 70% - 143% for CMAX – What is the impact on bioequivalence?, International Journal of Clinical Pharmacology and Therapeutics 39 (8): 350-355.
- Hauschke, D, Steinijans, VW (2000). The U.S. draft guidance regarding population and individual bioequivalence approaches: comments by a research-based pharmaceutical company, Statistics in Medicine 19: 2769-2774.
- Holmgren, EB (1999). Establishing equivalence by showing that a specified percentage of the effect of the active control over placebo is maintained, J. Biopharm. Stat 9(4): 651-659.
- Hung, JHM, Wang, S-J, and O'Neill, R (2005). A regulatory perspective on choice of margin and statistical inference issue in non-inferiority trials, Biometrical Journal 47 (1): 28-36.
- Hung, JHM, Wang, S-J, and O'Neill, R (2009). Challenges and regulatory experiences with non-inferiority trial design without placebo arm, Biometrical Journal 51 (2): 324-334.
- Inman, HF and Bradley, EL (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities, Communication in Statistics – Theory and Methods 18: 3851-3874.
- Inman, HF and Bradley, EL (1994). Hypothesis tests and confidence interval estimates for the overlap of two normal distributions with equal variances, Environmetrics 5: 167-189.
- Jensen, AL (1991). Separation coefficient for analysis of large environmental data sets: Ion concentrations of Lake Superior and Lake Huron. Environmetrics 2(3): 301-308.
- Koti, KM (2008). Use and abuse of statistics in design and analysis of active control trials, 2008 Proceedings of the Biopharmaceutical Section.
- Laster, LL, Johnson, MF (2003). Non-inferiority trials: the "at least as good as" criterion, Statistics in Medicine 22: 187-200.
- Lawrence, J (2005). Some remarks about the analysis of active control studies, Biometrical Journal 47: 616-622.

B&P BIOSTATISTICS & PROGRAMMING

❖ References

- Linacre, JM (1996). Overlapping normal distributions, Rasch Measurement Transactions 10 (1): 487-488.
- Mizuno, S, Yamaguchi, T, Fukushima, A, and Matsuyama,Y (2005). Overlap coefficient for assessing the similarity of pharmacokinetic data between ethnically different populations, Clinical Trials 2: 174-181.
- Mora, JO (1989). A new method for estimating a standardized prevalence of child malnutrition from anthropometric indicators, Bulletin of the World Health Organization 67: 133-142.
- Munk, A, Trampisch, HJ (2005). Therapeutic equivalence – Clinical issues and statistical methodology in non-inferiority trials, Biometrical Journal 47 (1): 7-9.
- Reiser, B and Faraggi, D (1999). Confidence intervals for the overlapping coefficient: The normal equal variance case, Department of Statistics, University of Haifa, Haifa 31999, Israel.
- Röhmel, J (1998). Therapeutic equivalence in investigations: statistical considerations, Statistics in Medicine 17: 1703-1714.
- Röhmel, J (2001). Statistical considerations of FDA and CPMP rules for the investigation of new anti-bacterial products. Statistics in Medicine 20: 2561-2571.
- Rom, DM and Hwang, E (1996). Testing for individual and population equivalence based on the proportion of similar responses, Statistics in Medicine 15: 1489-1505.
- Rothmann, M, Li, G, Chen, G, Chi, GYH, Temple, R and Tsou, HH (2003). Design and analysis of non-inferiority mortality trials in oncology, Statistics in Medicine 22: 239-264.
- Schall, R, Luus, HG (1993). On population and individual bioequivalence, Statistics in Medicine 12: 1109-1124.
- Senn, S (1993). Inherent difficulties with active control equivalence studies, Statistics in Medicine 12: 2367-2375.
- Senn, S (2001). Statistical issues in bioequivalence, Statistics in Medicine 20: 2785-2799.

B&P BIOSTATISTICS & PROGRAMMING

❖ References

- Senn, S (2005). 'Equivalence is different' – Some comments on therapeutic equivalence, Biometrical Journal 47 (1): 104-107.
- Stine, R and Heyse, JF (2001). Non-parametric estimates of overlap, Statistics in Medicine 20: 215-236.
- U.S. Food and Drug Administration (1992). Points to Consider. Division of Anti-infective Drug Products.
- U.S. Food and Drug Administration (1997). Antibacterial Drug Products: Use of Non-inferiority Studies to Support Approval.
- U.S. Food and Drug Administration (1998). Community-Acquired Pneumonia – Developing Antimicrobial Drugs for Treatment.
- U.S. Food and Drug Administration (2003). Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations.
- U.S. Food and Drug Administration (2008). Guidance for Industry on Diabetes Mellitus – Evaluating Cardiovascular Risk in New Anti-diabetic Therapies to Treat Type 2 Diabetes.
- U.S. Food and Drug Administration (2010). Guidance for Industry on Non-inferiority Clinical Trials, March 2010.
- Wiens, BL (2002). Choosing an equivalence limit for non-inferiority or equivalence studies, Controlled Clinical Trials 23: 2-14.