

What Equivalence, Non-Inferiority and Non-Superiority Mean

Daryl S. Paulson, PhD
BioScience Laboratories, Inc.

Traditional statistics are used in trying to find differences, say, between two products.

The hypotheses are:

$$H_0 : A = B$$

$$H_A : A \neq B \leftarrow \text{This is the hypothesis to be proved.}$$

Figure 1. Differences Between Products



Figure 1A shows that the two products are different, because the two confidence intervals do not overlap. Figure 1B shows that the products are the same, because the confidence intervals do overlap.

Equivalence

In the medical field, equivalence has become a goal. For example, a test product is very expensive and another product is inexpensive. The goal is to discover if they are equivalent, or the same, so the cheaper one can be sold.

In these tests, one is trying to determine if the two products (test and control) are the same within a given tolerance level δ . Out of pure error, you will find that one product has an advantage over the other, when there is not any difference. To protect for this, generally a small

tolerance level is added, which means the two products can differ within that amount of variation and still be termed equivalent.

Let us look at an example. The tolerance level is $\delta = 0.5 \log_{10}$. Because there is an upper and lower tolerance level in equivalency testing, they will be -0.5 and 0.5.

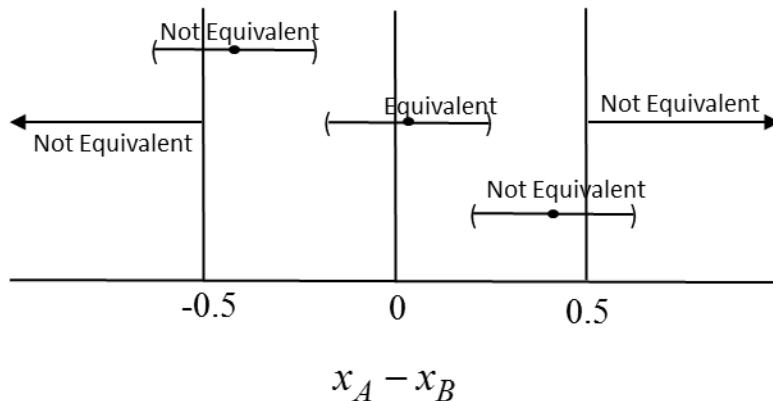
Two products are being evaluated: a Test Product (A) and a Control Product (B). If they are the same, then $\bar{x}_A - \bar{x}_B = 0$. But they can vary between -0.5 and 0.5 and still be termed equivalent. The hypotheses are:

$H_0 : x_A - x_B \leq -0.5$ and $H_0 : x_A - x_B \geq 0.5$ This is the hypothesis to be proved (non-equivalence).

$$H_A : -0.5 < x_A - x_B > 0.5 \text{ (equivalent)}$$

$$\delta = (-0.5, 0.5)$$

Figure 2. Equivalence Test

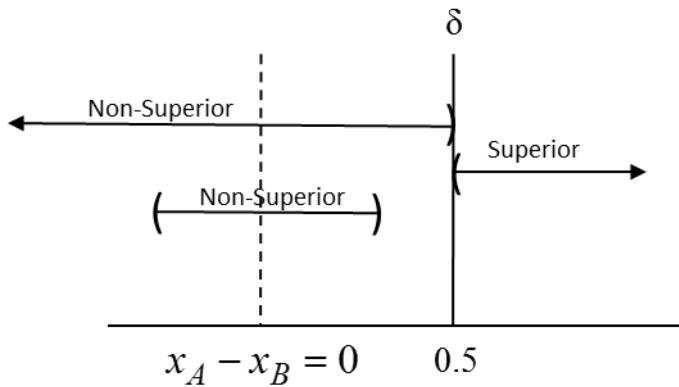


The test product is equivalent to the control product, as long as the 95% confidence interval is not below -0.5 or above 0.5.

Test Hypothesis for Superiority Test

Is the test product superior to the control product? Or is the test product non-superior to the control? The δ tells how much difference between $\bar{x}_A - \bar{x}_B$ you are willing to take in order to call it superior. Let us again use 0.5.

Figure 3. Non-Superiority Test



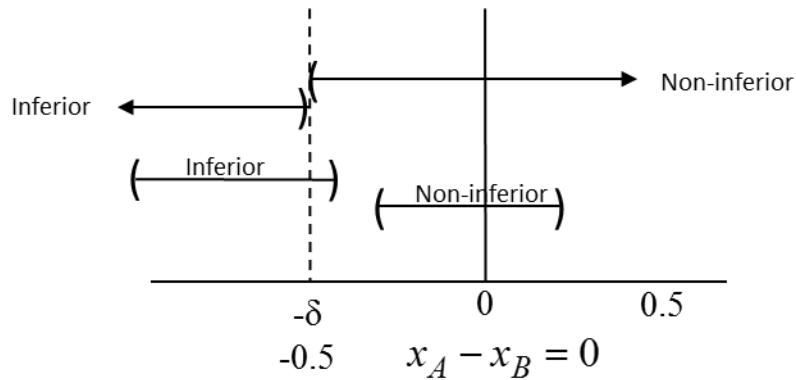
$H_0 : A - B \leq 0.5$ This is the hypothesis to be proved (non-superior).

$H_A : A - B > 0.5$ (superior)

Test Hypothesis for Inferiority Test

In this case, the test product must be above the control product by 0.5 or better.

Figure 4. Non-Inferiority Test



$$H_0 : A - B \leq -0.5 \quad \leftarrow \text{This is the hypothesis to be proved. (inferior)}$$

$$H_A : A - B > -0.5 \quad (\text{non-inferior})$$

The FDA's Supplemental Ruling to the Tentative Final Monograph

There are four components that must be addressed in evaluating a product, for example, a pre-operative skin preparation:

- Non-inferiority – the test product cannot be inferior to the control product;
- Superiority – test product must be superior to the vehicle product;
- At least 100 subjects per product must have the product applied for this study; and
- At the final sample time – six hours – all the reduction counts must be equal to or exceed the baseline value for both test and control products.

In these studies, three products must be used:

- Test product,
- Control product, and
- Vehicle product.

There is a margin of -0.5 that the FDA provides to the test versus control product. If your product is $-0.5 \log_{10}$ lower than the control, on the average, it is thought to be non-inferior. If it is -0.6, it is inferior.

Inferiority

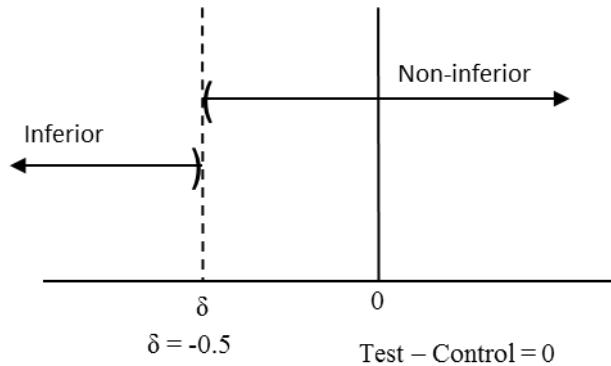
For Non-inferiority, the test product must be non-inferior to the control product with a margin of 0.5.

$$H_0 : A - B \geq -0.5 \text{ (non-inferior)}$$

$$H_A : A - B < -0.05 \text{ (inferior)}$$

A = Test, and B = Control

Figure 5. Non-Inferiority Test



Superiority

Test Products must be superior to the vehicle by:

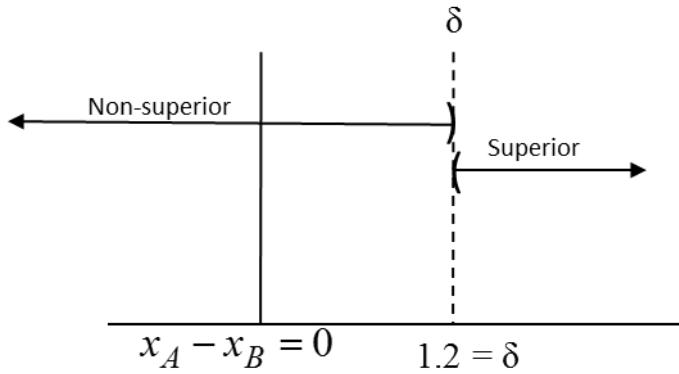
- δ [
- 1.2 \log_{10} pre-op skin prep at both inguinal and abdominal sites within 30 seconds or 10 minutes after the product has dried.
- 1.2 \log_{10} for a healthcare personnel handwash, after a single wash on each hand within 5 minutes
- 1.5 \log_{10} for a healthcare personnel handrub after a single wash on each hand within 5 minutes
- 0.5 \log_{10} for a surgical scrub after a single wash on each hand within 5 minutes

$$H_0 : A - B \leq \delta \text{ (non-superiority)} \leftarrow \text{This is the hypothesis to be proved.}$$

$$H_A : A - B > \delta \text{ (superiority)}$$

A = Test, and B = Control

Figure 6. Non-Superiority Test



This means that the test product minus the vehicle must be at least $1.2 \log_{10}$ better.

This can be a problem if your product has a very good delivery system and if the product is rather slow in working.

At least one hundred subjects have to be used for each product. This is not 300 subjects, because two products are applied to the two sides of a subject.

This can be determined by:

$$\frac{x(100)}{2} = \text{people to finish the study}$$

where:

$$x = \text{number of products} = 3$$

$$100 = \text{minimum number of subjects used}$$

$$\frac{3(100)}{2} = \frac{300}{2} = 150 \text{ subjects to finish the study, will give 300 sites for both}$$

inguinal and abdominal for the three products

The FDA uses a medical approach to the statistics. This means both types of error (alpha $\{\alpha\}$ and beta $\{\beta\}$) must be taking into account.

$$H_0 : A = B$$

$H_A : A \neq B$ ← This is the hypothesis to be proved.

Condition	Accept H_0	Reject H_0
H_0 true	Correct	Type I (α) $\alpha = 0.05$
H_0 false	Type II (β) $\beta = 0.20$	Correct

If $\alpha = 0.05$ and $\beta = 0.20$, Power = $1 - 0.20 = 0.80$ or 80%.

Confidence Level = $1 - 0.05 = 0.95$ or 95%.

We recommend performing a pilot study at our facility to:

- Determine if both the non-inferiority and superiority test work properly,
- Find the standard deviation of the three products, and
- Calculate the correct number of subjects to use.

One hundred percent (100%) of the subjects must be equal to or above the baseline counts.

Figure 7. Baseline Counts

