

**SHOULD THE NON-INFERIORITY MARGIN VARY WITH THE COMPARATOR
RATE?
AN ADAPTIVE STATISTICAL TEST**

Kem F. Phillips, Ph.D.

and

Michael Corrado, M.D.

Advanced Biologics LLC

24 Arnett Avenue

Suite 100

Lambertville, NJ 08530

Email: PHILLIPS@ADVBIOL.COM

Voice: 609-397-7891

Fax: 609-397-7892

Note: this is based on an article which will appear in **Statistics in
Medicine**[1].

Choosing the equivalence margin for a non-inferiority trial of an anti-infective medication involves both clinical and statistical considerations. This paper argues that for some trials the equivalence margin should depend on the underlying success rates, rather than being set in advance at a fixed level δ . A valid statistical test that adapts δ to the underlying success rates is described.

When the success rates for comparator drugs are well-established, a test of non-inferiority with a fixed δ is appropriate. However, a fixed δ may not be justifiable in a situation in which the expected success rates are difficult to estimate, such as for a new indication or for an unusual study design. As elucidated by Swartz[2], many organisms acquire resistance to anti-infective medications so that the success rates of these medications decrease over time. As an example, according to Swartz,

“Resistance to penicillin and penicillin-gentamicin synergism led in the late 1970s to widespread use of vancomycin in the treatment of life-threatening enterococcal infections. By the late 1980s vancomycin-resistant enterococci were reported, and in the mid 1990 these strains accounted for 13.6% of enterococcal isolates in intensive-care units in the United States.”

Swartz goes on to say that penicillin resistance in bacteria has increased to 20-25% in the United States. To combat resistance, medications may employ a new mechanism of action that is effective against pathogens that are resistant, or are soon to become resistant to approved drugs. These medications may be useful in treating diseases, especially in combination with other anti-infectives, even though their success rates are presently lower in comparison to other approved drugs. As microbial resistance to current drugs increases, the success rates for the newer drugs could become comparable. In addition to the problems posed by acquisition of resistance, anti-infective medications, such as vancomycin and other products currently in development, may target only some of the organisms that may infect a patient. This necessitates a more complicated study design than was used in the past, because other anti-infective medications must be

given for organisms that are not effectively treated by the test or comparator drugs. Finally, these powerful new medications are being increasingly used to treat recalcitrant and less well-understood indications such as bacteremia and neutropenia, which may not have a clear historical record on which to base predictions. All these factors may lead to lower, less predictable success rates for both the comparator and the test drugs. Statistical tests alone, especially tests with a fixed δ , cannot take these trends into account.

Until recently, the FDA's anti-infective "Points to Consider" guidance [3], which is no longer in effect, had been the basis for the statistical analysis of data from anti-infective trials. This guidance recommends using smaller δ s for higher observed success rates. Statistical problems with this procedure have been discussed by Röhmel[4] and others. But Röhmel points out that there are situations in which adapting δ to the underlying success rates is justified, and suggests that two criteria should be satisfied:

- “(1) There are good reasons (clinically and statistically) that the non-inferiority margin should vary with the response rate p of the standard drug or the better of the two.
- (2) The boundary curve of the equivalence margins should be smooth.”

In addition, as quoted by Röhmel, it was proposed by J. A. Lewis that the experimenter might:

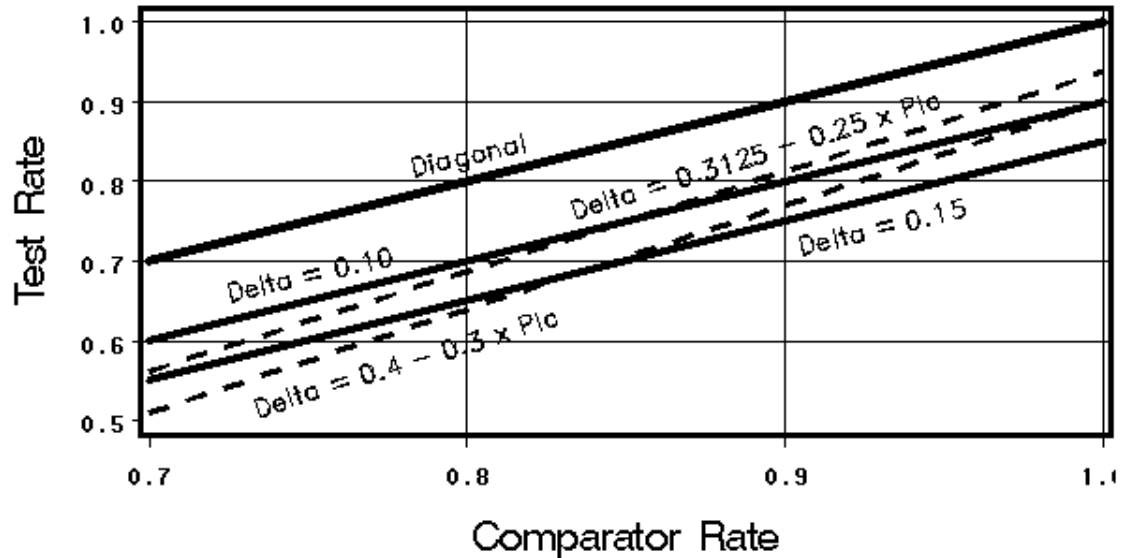
- “adopt the equivalence margin $\Delta(p)$ in such a way to the response rate p of the better of the two agents that the power of a study remains constant of a wide range of potential response rates, and is thus independent from the later observed response rates.”

A statistical test that fulfills these criteria can be developed as follows. In the standard test of non-inferiority the null statistical hypothesis is $\pi_t \leq \pi_c - \delta$, where π_t and π_c are the success rates in the test and comparator groups. The statistical test is invoked by computing a two-sided α -level confidence interval on $\pi_t - \pi_c$ and comparing the upper bound to $-\delta$. We may modify this simple test of

non-inferiority to allow δ to adapt to the comparator rate by incorporating the comparator rate into the expression for δ : $\delta(\pi_c) = \gamma + b\pi_c$. The null hypothesis becomes $\pi_t \leq \rho \pi_c - \gamma$, where $\rho = 1 - b$. The test statistic for H_0 can be developed in the same manner as the test statistic for H_{0F} . We reject H_0 if the one-sided α -level confidence bound on $\pi_t - (\rho\pi_c - \gamma)$ is greater than 0. This shows that the test can be interpreted as assessing whether the success rate of the test drug falls within the non-inferiority region. The mathematical formulas for the critical region, power, and sample size are modifications of the formulas for the standard test. The size of the test is near the nominal level and the power functions are smooth and increasing as we move away from the null hypothesis. A full exposition of this test and its properties are given by Phillips[1].

Many non-inferiority regions can be defined using this adaptive test, including regions that conform to the Lewis criterion. One interpretation of that proposal is that when the underlying success rates are equal, $\pi_t = \pi_c$, the power values should be constant over a reasonable range of values of the common success rate. This can be approximately achieved for certain combinations of ρ and γ . Figure 1, adapted from Röhmel, shows non-inferiority margins for four tests, two with fixed δ s, and two with values of ρ and γ which satisfy the Lewis criterion. The lines correspond to the boundaries of the inferiority/non-inferiority regions. On this graph $\pi_t = PIt$ and $\pi_c = P Ic$.

Figure 1: Non-Inferiority Margins for Four Tests



The sample sizes per group for 80% power at three success rates common to both test and comparator are shown in Table 1.

Table 1: Samples Sizes for Four Non-Inferiority Tests

	$\pi_S = \pi_T$		
	0.70	0.80	0.90
Fixed $\delta = 0.10$ ($\rho = 1, \gamma = 0.10$)	330	252	142
Adaptive, $\rho = 1.25, \gamma = 0.3125$ ($\delta = 0.3125 - 0.25 \pi_c$)	224	255	237
Fixed $\delta = 0.15$ ($\rho = 1, \gamma = 0.15$)	147	112	63
Adaptive, $\rho = 1.3, \gamma = 0.4$ ($\delta = 0.4 - 0.3 \pi_c$)	123	132	113

As expected, the sample size for tests with fixed δ s decreases with increasing success rates. The sample sizes for the adaptive tests, however, remain reasonably constant for success rates between 0.70 and 0.90.

In conclusion, the choice of δ , or non-inferiority margin, is somewhat arbitrary. Strict adherence to a statistical standard in approving anti-infective medications would exclude many factors of great medical importance, especially the increase in microbial resistance. A drug that narrowly misses a statistical target but is effective against a new strain of pathogen may still be useful in medical practice. An adaptive- δ test allows a greater range of alternatives in setting up appropriate statistical hypotheses, and can be a part of a process that couples correct statistical procedures with clinical acumen and judgement.

REFERENCES

1. Phillips K. A New Test of Non-inferiority for Anti-Infective Trials. To be published in *Statistics in Medicine*.
2. Swartz MN. Impact of *Antimicrobial Agents and Chemotherapy* from 1972 to 1998. *Antimicrobial Agents and Chemotherapy* 2000; 44: 2009-2016.
3. Röhmel J. Therapeutic equivalence investigations: statistical considerations, *Statistics in Medicine* 1998; 17:1703-1714.
4. U.S. Food and Drug Administration. Points to consider, Division of Anti-infective Drug Products, Clinical Development and Labeling of Anti-Infective Drug Products. 1992.