

# On the Analysis of a Combination of Two Primary Efficacy Measures

Karl E. Peace<sup>1</sup>

## Abstract

In many clinical trials of drugs, there are two primary efficacy measures. In such studies, the traditional framework for analysis, is to construct two separate univariate null and alternative hypotheses and test each of these at some appropriate Type I error level. Under this framework, it is desirable to reject each null hypothesis ( $H_{01}, H_{02}$ ); i.e. 'to win on both.' To 'win on both' is a rather stringent requirement of statistical efficacy for a test drug. For if 'winning on both' is interpreted as the statistical rejection of each null hypothesis, each at the 0.05 level of significance, then 'winning on both', in the case of independence of the two efficacy measures, has an associated 0.0025 level of significance. Therefore, a method which utilizes information on both response measures, but which does not exact this severity, is desirable.

As an alternative to the traditional approach, we consider that clinical response (CR) has been defined on each of two efficacy measures separately, so that two dichotomous random variables,  $X_1$  and  $X_2$  are produced. We then define the primary efficacy measure ( $Y$ ) to be the sum of  $X_1$  and  $X_2$ . In this approach, one is interested in the individual patient 'winning on both' – perhaps a more stringent requirement of clinical effectiveness of a test drug. In addition, the multiple endpoint problem is avoided, and consequently, what Type I error rate, should be used to test each individual null hypothesis. One would simply test,  $H_{03} : P_{3B} = P_{3A}$  versus  $H_{a3} : P_{3B} > P_{3A}$ , where  $P_{3A}$  and  $P_{3B}$  denote the true proportions of patients who would respond according to the clinical response criteria for both efficacy variables, when treated with A and B respectively.

The methods are applied to data from a clinical trial comparing a drug to placebo, in the treatment of patients suffering from Alzheimer's disease. The two primary efficacy measures in the study were clinical global impression of change (CGIC) ratings and ratings according to the cognitive component of the Alzheimer's disease assessment scale (ADAS-COG).

## I. INTRODUCTION

In many placebo-controlled clinical trials of new drugs, there are two primary efficacy measures. For example in the development of a fixed combination drug for the treatment of allergic rhinitis, one efficacy measure is nasal airway resistance (reflecting the degree of congestion), whereas another is relief of hay fever-like symptoms [1]. In the development of anti-ulcer drugs, such as the H<sub>2</sub>-receptor antagonists, one efficacy measure is the proportion of patients whose ulcers heal, whereas another is the proportion of patients free of pain.[2]

In such studies, the traditional framework for analysis, is to construct two separate univariate null and alternative hypotheses and test each of these at some appropriate Type I error level. The separate univariate hypotheses are:

---

<sup>1</sup> FASA, GCC Distinguished Cancer Scholar, Senior Research Scientist and Professor of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, PO Box 8148-01, Statesboro, GA USA 30460. Email address: [kepeace@georgiasouthern.edu](mailto:kepeace@georgiasouthern.edu), [peacekarl@cs.com](mailto:peacekarl@cs.com)

$$\begin{aligned} H_{01} : P_{1B} = P_{1A} \text{ versus } H_{a1} : P_{1B} > P_{1A}, \\ H_{02} : P_{2B} = P_{2A} \text{ versus } H_{a2} : P_{2B} > P_{2A}, \end{aligned}$$

where A reflects a placebo group, B reflects the new drug group,  $P_{1A}$  and  $P_{1B}$  reflect the true mean effects of A and B in terms of the first efficacy measure, respectively; and  $P_{2A}$  and  $P_{2B}$  reflect the true mean effects of A and B in terms of the second efficacy measure, respectively.

From a drug development point of view, it is desirable to reject each of  $H_{01}$  and  $H_{02}$ ; *i.e.*, 'to win on both.' To 'win on both' is a rather stringent requirement of statistical efficacy for a new drug. For if 'winning on both' is interpreted as the statistical rejection of  $H_{01}$  and  $H_{02}$ , each at the 0.05 level of significance, then 'winning on both', in the case of independence of the two efficacy measures, has an associated 0.0025 level of significance. Therefore, a method which utilizes information on both response measures, but which does not exact this severity, would be desirable.

## II. ALTERNATIVE APPROACH

From the patient's point of view, and from the practicing clinicians point of view, it is desirable that the patient 'wins on both' efficacy measures. In the case of clinical development of H<sub>2</sub>-receptor antagonists as anti-ulcer agents, patients are screened for admission into an anti-ulcer clinical trial on the basis of epigastric, ulcer-like pain. Such patients are subjected to an endoscopy, and if this reveals the presence of an ulcer, they are allowed to enter the trial. Therefore, such patients have both pain and ulcers present when treatment with the new drug or placebo begins. The clinical therapeutic goal is to heal the ulcer and alleviate the pain in each patient, rather than healing the ulcer in one subset of patients and alleviating pain in another subset of patients.

As an alternative to the traditional approach, we consider that clinical response has been defined on each of two efficacy measures separately, so that two dichotomous random variables are produced. That is, let  $X_1$  denote one efficacy variable, and  $X_2$  denote the other, where on an individual patient basis,  $X_i = 1$  if the definition of clinical response holds, and  $X_i = 0$  if the definition of clinical response does not hold;  $i = 1, 2$ .

We then define the primary efficacy measure ( $Y$ ) to be the event that a patient clinically responds according to both variables. Peace([1],[2]) has used this approach successfully, in a number of studies. Often when one 'wins' on one univariate hypothesis, and does not on the other, one 'wins' in terms of the proportion of patients who respond on both measurement scales.

As previously indicated, in this latter approach, one is interested in the individual patient 'winning on both' – perhaps a more stringent requirement of clinical effectiveness of a new drug. In addition, the statistics are cleaner, since one does not have to worry about what Type I error rate should be used to test each individual null hypothesis. One simply tests the null hypothesis  $H_{03}$  versus  $H_{a3}$  at 0.05 level of significance, where  $H_{03}$  and  $H_{a3}$  are specified as:

$$H_{03} : P_{3B} = P_{3A} \text{ versus } H_{a3} : P_{3B} > P_{3A},$$

where  $P_{3A}$  and  $P_{3B}$  denote the true proportions of patients who would respond according to the clinical response criteria for both efficacy variables, when treated with A and B, respectively.

### III THE JOINT DISTRIBUTION OF $X_1$ AND $X_2$

#### A. Bivariate Bernoulli Distribution

In the setting described in Section II,  $X_1$  and  $X_2$  have a joint (bivariate) Bernoulli or point binomial distribution.

This distribution,  $P(X_1 = x_1; X_2 = x_2)$ , is given by:

$$P(x_1; x_2) = (P_1^{x_1})(P_2^{x_2})(Q_1^{1-x_1})(Q_2^{1-x_2})(1 + D \cdot F), \tag{1}$$

where  $D = R\sqrt{P_1Q_1P_2Q_2},$  (2)

$$F = (x_1 - P_1)(x_2 - P_2) / P_1Q_1P_2Q_2, \tag{3}$$

$$Q_i = 1 - P_i, \tag{4}$$

$$X_i = 0 \text{ or } 1, \tag{5}$$

and  $R$  is the true correlation coefficient between  $X_1$  and  $X_2$ , for  $i = 1, 2$ .

#### B. Distribution Properties – Joint Bernoulli Distribution

It is easy to show that the univariate or marginal distribution of  $X_i; i = 1, 2$ , given that  $X_1$  and  $X_2$  have a joint bivariate Bernoulli distribution,  $P(x_i)$ , given by:

$$P(x_i) = P(X_i = x_i) = P_i^{x_i} Q_i^{1-x_i} \tag{6}$$

with mean  $E(X_i)$ , and variance,  $V(X_i)$ , given by:

$$E(X_i) = P_i \tag{7}$$

and

$$V(X_i) = P_i(1-P_i) \tag{8}$$

where  $X_i = 0, 1$ , for  $i = 1, 2$ .

Further, it is easy to show that  $D$  is the covariance of  $X_1$  and  $X_2$ ; *i.e.*

$$D = \text{Cov}(X_1, X_2) \tag{9}$$

Likewise, it is easy to show that the conditional distribution of  $X_2$ (say) given that  $X_1 = x_1$ , where  $X_1$  and  $X_2$  have a joint bivariate Bernoulli distribution given by:

$$P(X_2 | X_1 = x_1) = P_2^{x_2} Q_2^{1-x_2} (1 + D \cdot F) \quad (10)$$

The mean,  $E(X_2 | X_1 = x_1)$  and variance,  $V(X_2 | X_1 = x_1)$  of this conditional distribution are given by:

$$E(X_2 | X_1 = x_1) = P_2 + R\sqrt{P_2 Q_2 P_1 Q_1} (x_1 - P_1) \quad (11)$$

and,

$$V(X_2 | X_1 = x_1) = E(X_2 | X_1 = x_1)[1 - E(X_2 | X_1 = x_1)] \quad (12)$$

### C. Maximum Likelihood Estimators – Joint Bernoulli.

Distribution Parameters:

Suppose the two-by-two Table 1, represents a random sample from the joint, bivariate, Bernoulli distribution. From maximum likelihood theory, it is easy to show that the maximum likelihood estimators of the parameters of the distribution are given by:

$$\hat{P}_1 = N_{.1} / N_{..}, \hat{Q}_1 = 1 - \hat{P}_1 \quad (13)$$

$$\hat{P}_2 = N_{1.} / N_{..}, \hat{Q}_2 = 1 - \hat{P}_2 \quad (14)$$

$$\hat{R} = (N_{00} / N_{11} - N_{01} / N_{10}) / \sqrt{N_{0.} N_{1.} N_{.0} N_{.1}} \quad (15)$$

and

$$\hat{D} = \hat{R} \sqrt{\hat{P}_1 \hat{Q}_1 \hat{P}_2 \hat{Q}_2} \quad (16)$$

**Table 1:** Joint Bernoulli Random Sample

		X <sub>1</sub>		
		0	1	
X <sub>2</sub>	0	N <sub>00</sub>	N <sub>01</sub>	N <sub>0.</sub>
	1	N <sub>10</sub>	N <sub>11</sub>	N <sub>1.</sub>
		N <sub>.0</sub>	N <sub>.1</sub>	N <sub>..</sub>

It should be noted that  $\hat{R}$  is the same measure of correlation as Yule's ([3][4])  $V$ . Further, it should be noted that

$$\chi^2 = N_{..} \hat{R}^2. \quad (17)$$

Thus,  $\hat{R} = \sqrt{\chi^2 / (N_{..})}$  where  $\chi^2$  is the Pearson's chi-square computed from Table 1.

**IV. THE UNIVARIATE DISTRIBUTION OF  $Y = X_1 + X_2$**

The random variable  $Y$ , defined as the sum of  $X_1$  and  $X_2$ , takes on the distinct values 0, 1, and 2. It has the value 0 when both  $X_1$  and  $X_2$  are 0. It has the value 1 when either  $X_1$  is 0 and  $X_2$  is 1, or when  $X_1$  is 1 and  $X_2$  is 0. It has the value 2 when both  $X_1$  and  $X_2$  are 1.

**A. The Point Univariate Trinomial Distribution**

The univariate distribution of  $Y$ , is the point trinomial distribution,  $P(y)$ , given by equations (19), (20) and (21):

$$P(Y = 0) = P(X_1 = 0; X_2 = 0) = Q_1Q_2 + D, \tag{19}$$

$$\begin{aligned} P(Y = 1) &= P(X_1 = 0; X_2 = 1) + P(X_1 = 1; X_2 = 0) \\ &= P_1Q_2 + P_2Q_1 - 2D, \end{aligned} \tag{20}$$

$$P(Y = 2) = P(X_1 = 1; X_2 = 1) = P_1P_2 + D. \tag{21}$$

**B. Distributional Properties – Point Trinomial Distribution**

Using the statistical definitions of the mean and variance of a distribution, it is easy to show that the mean  $E(Y)$  and variance  $V(Y)$  of the univariate, point trinomial distribution, are given by:

$$E(Y) = P_1 + P_2 \tag{22}$$

$$V(Y) = P_1 Q_1 + P_2 Q_2 + 2D \tag{23}$$

**C. Maximum Likelihood Estimators – Point Trinomial Distribution Probabilities**

By applying maximum likelihood theory to the data in Table 1, it is easy to show that the maximum likelihood estimators of the point trinomial probabilities:  $P(Y = 0)$ ,  $P(Y = 1)$ , and  $P(Y = 2)$ , are given by:

$$P(Y = 0) = N_{00} / N_{..}, \tag{24}$$

$$P(Y = 1) = N_{01} + N_{10} / N, \tag{25}$$

and,

$$P(Y = 2) = N_{11} / N_{..} \tag{26}$$

It is reassuring to know that when the maximum likelihood estimators given in equations (13), (14), (15), and (16), are substituted into equations (19), (20), and (21), the

latter equations reduce to equations (24), (25), and (26). That this is so is shown below for equation (26).

First of all note that  $D$  is estimated as:

$$\hat{D} = \frac{N_{00}N_{11} - N_{01}N_{10}}{\sqrt{N_{00}N_{11}N_{01}N_{10}}} \cdot \frac{\sqrt{N_{.1}N_{.0}N_{1.}N_{0.}}}{\sqrt{N_{..} \cdot N_{..} \cdot N_{..} \cdot N_{..}}} = (N_{00}N_{11} - N_{01}N_{10}) / N_{..}^2 \quad (27)$$

Now,

$$\begin{aligned} P_1P_2 + D &= (N_{.1} / N_{..})(N_{.1} / N_{..}) + (N_{00}N_{11} - N_{01}N_{10}) / N_{..}^2 \\ &= [(N_{01} / N_{11})(N_{10} / N_{11}) + (N_{00}N_{11} - N_{01}N_{10})] / N_{..}^2 \\ &= N_{11}(N_{01} + N_{10} + N_{11} + N_{00}) / N_{..}^2 \\ &= N_{11}N_{..} / N_{..}^2 = N_{11} / N_{..} \end{aligned} \quad (28)$$

#### D. Probability Calculations for the Trinomial Distribution

Suppose that  $N_{..}$  is the size of a random sample (Table 1) from the point trinomial distribution. Further let  $N_{00}$  denote the number of 0's,  $N_{11}$  denote the number of 2's, and  $N_1 = N_{..} - N_{00} - N_{11}$  denote the number of 1's in the sample. Then the probability,  $P(N_{00}; N_1; N_{11})$ , associated with the event that  $Y = 0$   $N_{00}$  times, and  $Y = 1$   $N_1$  times, and  $Y = 2$   $N_{11}$  times, is given by:

$$P(N_{00}; N_1; N_{11}) = (N_{..}! / N_{00}!N_1!N_{11}!) [P(Y = 0)N_{00}] [P(Y = 1)N_1] [P(Y = 2)N_{11}] \quad (29)$$

where the individual probabilities:  $P(Y = 0)$ ,  $P(Y = 1)$ , and  $P(Y = 2)$ , are given respectively by equations (19), (20) and (21).

The probability given in equation (29) may be estimated by substituting into the equation, the maximum likelihood estimates given by equations (24), (25), and (26).

#### V. CONTRASTING HYPOTHESES

The results given in Sections III and IV, were without reference to a particular treatment group. In this Section, we assume that the results of Sections III and IV apply individually to a new drug group ( $B$ ) and to a placebo group ( $A$ ). In addition to the parameters of the distributions being indexed by 1 or 2 (for the two primary efficacy variables,  $X_1$  and  $X_2$ ), they will also be indexed by the treatment group labels ( $A, B$ ).

As discussed in Section I, the separate univariate hypotheses, corresponding to  $X_1$  and  $X_2$  are given by:

$$H_{01} : P_{1B} = P_{1A} \text{ versus } H_{a1} : P_{1B} > P_{1A},$$

$$H_{02} : P_{2B} = P_{2A} \text{ versus } H_{a2} : P_{2B} > P_{2A}$$

Also as discussed in Section II, the univariate hypotheses, corresponding to  $Y = X_1 + X_2$  are given by:

$$H_{03} : P_{3B} = P_{3A} \text{ versus } H_{a3} : P_{3B} > P_{3A},$$

where  $P_{3A}$  and  $P_{3B}$  denote the true proportions of patients who would respond according to the clinical response criteria for both efficacy variables, when treated with  $A$  and  $B$ , respectively.

From section IV, we see that  $H_{03}$  and  $H_{a3}$  may be written as:

$$H_{03} : (P_{1B} P_{2B}) + D_B = (P_{1A} P_{2A}) + D_A$$

$$H_{a3} : (P_{1B} P_{2B}) + D_B > (P_{1A} P_{2A}) + D_A$$

and therefore,

$$P_{3B} = (P_{1B} P_{2B}) + D_B, \tag{30}$$

and

$$P_{3A} = (P_{1A} P_{2A}) + D_A. \tag{31}$$

So we see, for example, that the true proportion of patients who would respond according to the clinical response criteria for both efficacy variables  $X_1$  and  $X_2$ , when treated with  $B$ , is the product of the true proportion of patients who would respond according to the clinical response criteria for  $X_1$ , and the true proportion of patients who would respond according to the clinical response criteria for  $X_2$ , plus the covariance between  $X_1$  and  $X_2$ .

Now,  $P_{1B}$  and  $P_{2B}$  (or  $P_{1A}$  and  $P_{2A}$ ) will usually be less than 1. Therefore  $P_{1B} P_{2B}$  will be less than  $P_{1B}$  or  $P_{2B}$ , individually. Thus, in the case of no or weak correlation between  $X_1$  and  $X_2$  in treatment group  $B$  (or in treatment group  $A$ ),  $P_{3B}$  (or  $P_{3A}$ ) would be expected to be less than either  $P_{1B}$  or  $P_{2B}$  (or  $P_{1A}$  or  $P_{2A}$ ). However, for strongly correlated variables,  $P_{3B}$  (or  $P_{3A}$ ) may be equal to or exceed  $P_{1B}$  or  $P_{2B}$  (or  $P_{1A}$  or  $P_{2A}$ ). So the degree to which both  $P_{1B} P_{2B}$  is greater than  $P_{1A} P_{2A}$ , and/or  $D_B$  is greater than  $D_A$ , is consistent with  $H_{a3}$ : ( $P_{3B} > P_{3A}$ ).

It is worth noting that  $E(X_{iA}) = P_{iA}$ , and  $E(X_{iB}) = P_{iB}$ ,  $i = 1, 2$ ; where  $A$  and  $B$  denote two treatment groups, and where  $E$  denotes expected value. And therefore, both  $H_{01}$  and  $H_{02}$  reflect comparing the treatment groups  $A$  and  $B$ , in terms of the centers of the distributions of the  $X_i$ . This is not the case for  $H_{03}$ . Under  $H_{03}$ , the treatment groups are being compared in terms of the true portions of patients who fall into the multinomial category  $Y = 2$  – which is the same as the proportions of patients ‘who win’ on both efficacy variables. As has been mentioned before, this is the same as the proportion of patients who fall into the (1,1) category in the joint bivariate distribution of  $X_1$  and  $X_2$ .

## VI. TESTS OF HYPOTHESES

There are three experimental situations which lead to frequency data being expressed in the form of two-by-two contingency tables. The first is when both margins are fixed. The

second is when one margin is fixed. The last is when neither margin is fixed. These situations are referred to as two-by-two independence trials, two-by-two comparative trials, and two-by-two double-dichotomy trials, respectively. It is well known that when both margins are fixed, Fisher's exact test [7], is the uniformly most powerful unbiased UMPU test of association between  $X_1$  and  $X_2$ . However, Tocher [8] has shown the remarkable result, that Fisher's exact test is also the UMPU test when one or no margin is fixed. Parenthetically, it should be pointed out that discreteness of the distributions limits the size or Type I error of the tests. And thus, these results strictly hold only if randomization is permitted so that tests of any size may be obtained. Therefore, Fisher's exact test may be used to test  $H_{o1}$ ,  $H_{o2}$  and  $H_{o3}$ , respectively.

To illustrate the two-by-two tables which would be used in carrying out the tests, Tables 2 and 3 below, provide frequencies representing the  $X_1$ -by- $X_2$  classifications for treatment groups  $A$  and  $B$ , respectively:

Table 2: Joint Bernoulli Random Sample From Treatment A Population

	$X_{A1}$			
	0	1		
$X_{A2}$	0	$N_{A00}$	$N_{A01}$	$N_{A0\cdot}$
	1	$N_{A10}$	$N_{A11}$	$N_{A1\cdot}$
		$N_{A\cdot 0}$	$N_{A\cdot 1}$	$N_{A\cdot\cdot}$

Table 3: Joint Bernoulli Random Sample From Treatment B Population

	$X_{B1}$			
	0	1		
$X_{B2}$	0	$N_{B00}$	$N_{B01}$	$N_{B0\cdot}$
	1	$N_{B10}$	$N_{B11}$	$N_{B1\cdot}$
		$N_{B\cdot 0}$	$N_{B\cdot 1}$	$N_{B\cdot\cdot}$

From Tables 2 and 3, the two-by-two tables for testing  $H_{o1}$  and  $H_{o2}$ , appear below as Tables 4 and 5, respectively:

Table 4: Frequencies for Testing  $H_{o1}$

Group	0	1	
A	$N_{A\cdot 0}$	$N_{A\cdot 1}$	$N_{A\cdot\cdot}$
B	$N_{B\cdot 0}$	$N_{B\cdot 1}$	$N_{B\cdot\cdot}$
Total	$N_{\cdot 0}$	$N_{\cdot 1}$	$N_{\cdot\cdot}$

Table 5: Frequencies for Testing  $H_{o2}$

Group	0	1	
A	$N_{A0\cdot}$	$N_{A1\cdot}$	$N_{A\cdot\cdot}$
B	$N_{B0\cdot}$	$N_{B1\cdot}$	$N_{B\cdot\cdot}$
Total	$N_{0\cdot}$	$N_{1\cdot}$	$N_{\cdot\cdot}$



Finally, from Tables 2 and 3, the two-by-two table for testing  $H_{03}$ , appears below as Table 6:

**Table 6:** Frequencies for Testing  $H_{03}$

Group	0 or 1	2	
A	$N_{A.} - N_{A11}$	$N_{A11}$	$N_{A..}$
B	$N_{B.} - N_{B11}$	$N_{B11}$	$N_{B..}$
Total	$N_{...} - N_{.11}$	$N_{.11}$	$N_{...}$

**VII. EXAMPLE**

Data are given in Tables 7 and 8 from a clinical trial comparing a drug to placebo in the treatment of patients suffering from Alzheimer’s disease. The two primary efficacy measures in the study were clinical global impression of change (CGIC) ratings, and ratings according to the cognitive component of the Alzheimer’s disease assessment scale (ADAS-COG). The clinical response criterion ( $X_1$ ) for CGIC represented no deterioration from baseline by the sixth week of the study. The clinical response criterion ( $X_2$ ) for ADAS-COG represented at least a two-point improvement from baseline by the end of the study. Data reflecting  $X_1$  and  $X_2$  for each treatment group are presented below in Tables 7 and 8:

**Table 7:** Joint Bernoulli Random Sample From Placebo. Group (A)

		$X_{A1}$		
		0	1	
$X_{A2}$	0	21	42	63
	1	0	9	9
		21	51	72

**Table 8:** Joint Bernoulli Random Sample From Drug Group (B)

		$X_{B1}$		
		0	1	
$X_{B2}$	0	14	30	44
	1	2	25	27
		16	55	71

From Tables 7 and 8, the two-by-two tables for testing  $H_{01}$  and  $H_{02}$ , appear below as Tables 9 and 10, respectively:

**Table 9:** Frequencies for Testing  $H_{01}$

Group	0	1	
A	21	51	72
B	16	55	71
Total	43	106	413

Table 10: Frequencies for Testing Ho2

Group	0	1	
A	63	9	72
B	44	27	71
Total	107	36	143

Finally, from Tables 7 and 8, the two-by-two table for testing  $H_{03}$ , appears below as Table 11:

Table 11: Frequencies for Testing Ho3

Group	0 or 1	2	
A	63	9	72
B	46	25	71
Total	109	34	143

For the placebo group ( $A$ ), one obtains the estimates:

$$R_A = 0.243, D_A = 0.0364, P_{A1} = 0.708, P_{A2} = 0.125$$

For the drug group ( $B$ ), one obtains the estimates:

$$R_B = 0.284, D_B = 0.0576, P_{B1} = 0.775, P_{B2} = 0.380$$

Therefore, as estimate of the difference between  $B$  and  $A$ , in terms of the proportion of patients 'winning' on both  $X_1$  and  $X_2$  is given by:

$$\begin{aligned} &= (0.775)(0.380) + 0.0576 - (0.708)(0.125) - 0.0364 \\ &= 0.2271 \end{aligned}$$

which has an associated one-tailed Fisher's exact P-value of 0.001. Therefore these data reflect a statistically significant benefit of drug as compared with placebo.

## VIII. SUMMARY

In clinical trials of drugs where there are two primary efficacy measures, the traditional framework for analysis is to construct two separate univariate null and alternative hypotheses and test each of these at some appropriate Type I error level. The separate univariate hypotheses are:  $H_{01} : P_{1B} = P_{1A}$  versus  $H_{a1} : P_{1B} > P_{1A}$ , and  $H_{02} : P_{2B} = P_{2A}$  versus  $H_{a2} : P_{2B} > P_{2A}$ , where  $A$  reflects a control group,  $B$  reflects a test drug group,  $P_{1A}$  and  $P_{1B}$  reflect the true mean effects of  $A$  and  $B$  in terms of the first efficacy measure, respectively; and  $P_{2A}$  and  $P_{2B}$  reflect the true mean effects of  $A$  and  $B$  in terms of the second efficacy measure, respectively.

Under this framework, it is desirable to reject each of  $H_{01}$  and  $H_{02}$ ; *i.e.* 'to win on both' To 'win on both' is a rather stringent requirement of statistical efficacy for a test drug.

For if 'winning on both' is interpreted as the statistical rejection of  $H_{01}$  and  $H_{02}$ , each at the 0.05 level of significance, then 'winning on both', in the case of independence of the two efficacy measures, has an associated 0.0025 level of significance. Therefore, a method which utilizes information on both response measures, but which does not exact this severity, is desirable.

We've proposed an alternative to the traditional approach. First clinical response ( $CR$ ) is defined on each of two efficacy measures separately, so that two dichotomous random variables,  $X_1$  and  $X_2$ , are produced. We then define the primary efficacy measure ( $Y$ ) to be the sum of  $X_1$  and  $X_2$ . In this approach, we are interested in the individual patient 'winning on both.' This is perhaps a more stringent requirement of clinical effectiveness of a test drug. In addition, one doesn't have to worry about the multiple endpoint problem, and consequently, what Type I error rate, should be used to test each individual null hypothesis. One simply tests, using Fisher's exact test,  $H_{03} : P_{3B} = P_{3A}$  versus  $H_{a3} : P_{3B} > P_{3A}$ , where  $P_{3A}$  and  $P_{3B}$  denote the true proportions of patients who would respond according to the clinical response criteria for both efficacy variables, when treated with  $A$  and  $B$ , respectively.

Since the  $X_i, i = 1, 2$ , are indicator variables of  $CR$  ( $X_i = 1$  if  $CR$  and  $X_i = 0$  if not  $CR$ ), the joint distribution of  $X_1$  and  $X_2$  is a Bernoulli distribution, with marginal probabilities of  $CR$  being  $P_1$  and  $P_2$ , and covariance  $D$ . Further, the distribution of  $Y$  is a point trinomial distribution, with  $P_{3B} = P_{1B} P_{2B} + D_B$ , and  $P_{3A} = P_{1A} P_{2A} + D_A$ . So that,  $H_{a3}$  is consistent with  $P_{1B} P_{2B} > P_{1A} P_{2A}$  and/or  $D_B > D_A$ . Also, the maximum likelihood estimators of  $P_{3B}$  and  $P_{3A}$  are the proportions of patients in treatment groups  $B$  and  $m$ , respectively, who win on both  $X_1$  and  $X_2$ . The methods were applied to data from a clinical trial compared to placebo in the treatment of patients suffering from Alzheimer's disease. The two primary efficacy measures in the study were clinical global impression of change (CGIC) ratings and ratings according to the cognitive component of the Alzheimer's disease assessment scale (ADAS-COG). The clinical response criterion ( $X_1$ ) for CGIC represented no deterioration from baseline by the end of the study. The clinical response criterion ( $X_2$ ) for ADAS-COG represented at least a two-point improvement from baseline by the end of the study. Based upon analyses of these data using the method proposed, the drug was statistically significantly ( $P = 0.001$ ) more effective than placebo. This means that the proportion of patients responding according to both CGIC and ADAS-COG clinical response criteria is significantly greater in the drug group, than in the placebo plus lecithin group. By contrast, the P-values for CGIC and ADAS-COG responders individually, are 0.238 and 0.0004, respectively.

### References

- [1] Diamond L, Gerson K, Cato a, Peace KE, Perkins JG: "An Evaluation of Triprololodine and Psuedoephedrine in the Treatment of Allergic Rhinitis," *Annals of Allergy*. June, 1981.
- [2] Peace KE etal: "A Single Nocturnal dose of Cimetidine in Active Duodenal Ulcer: Statistical Considerations in the Design, Analysis, and Interpretation of a Clinical Trial." *Post Graduate Medicine*. Nov., 1985.
- [3] Yule GU: "On Association of Attributes in Statistics." *Phil. Trans.*, a, 194: 257, 1900.
- [4] Yule GU: "On Methods of Measuring the Association of Two Attributes." *J.R. Statist. Soc.*, 75: 579, 1912.
- [5] Pearson K: "On Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be Reasonably supposed to have arisen in Random Sampling." *Phil. Mag.*, (5)50: 157, 1900.
- [6] Kendall MG, Stuart A: *The Advanced Theory of Statistics*. Hafner Publishing Company, New York, 2: 571, 1973.
- [7] Fisher RA: *The Design of Experiments*. Oliver and Boyd, Edinburgh, 1935.
- [8] Tocher KD: "Extension of the Neyman-Pearson Theory of Tests to Discontinuous Variates." *Biometrika*, 37: 130, 1950.