Define a more powerful endpoint in a longitudinal trial with information of correlation coefficient

Ruji Yao[1], Qing Li[1], Wen-Chi Wu[1], Norman Y. Yao[2]
[1]Merck & Co., Inc., NJ, USA
[2]Department of Physics, UC Berkeley, CA, USA

Abstract
For a longitudinal clinical trial, an efficacy endpoint can be defined on a single time point at the end of trial; or an average of data over all post baseline time points, like in an allergy study; or partial post baseline time points, like in an asthma study. When using multiple time points to define an efficacy endpoint, it is usually assumed that the treatment effect is close to the expected trial value over the interval used. Here, we will define a formula to diagnose how close the expected and actual treatment effect must be in order to realize an equal or better power when comparing to a single time endpoint. Of course, it is well known that an endpoint defined using multiple time points is less sensitive to missing values and we will extend our formula to account for a given fraction of missing values.

1. Introduction

Let us consider two example case asthma studies:

- Study 1. total 4 visits: baseline, Week 2, 4 and 6. The primary efficacy endpoint for each subject was the average change from Baseline of 2 time points, Week 4 and 6.

- Study 2. total 7 visits: baseline, Week 2, 4 , 6, 8,10 and 12. The primary efficacy endpoint for each subject was the average change from Baseline of 4 time points, Week 6, 8, 10 and 12.

It is obvious that the endpoint could be defined in many different ways by using either a single or multiple time points.

The key question we will try to address here is: What is the precise criteria to define a "better" endpoint?

2. Consider a design of 2 treatment groups (active vs. placebo) with K post baseline time points.

Assumption:
Change from baseline at time K are $\mu_p \; and \; \mu_a$, respectively, for placebo and active groups. $n = n_a = n_p$ is sample size for each group.

We use the t-statistics to compare treatment difference at time K

$$T = \frac{\delta}{\sigma\sqrt{\frac{1}{n_a} + \frac{1}{n_p}}}$$

Where treatment effect $\delta = \mu_a - \mu_p$ and σ is the pooled standard deviation.

3. Consider an endpoint using the average of time points K-1 and K.

More assumptions:
Treatment effect is $p\delta$ at time K-1.
If $p = 0.9$, then the treatment effect at time K-1 is assumed as 90% of time K.

Change from baseline at time K-1 are now $p\mu_p \; and \; p\mu_a$, respectively, for placebo and active groups.

Variance matrix of time K-1 and K $\begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}$ (1)

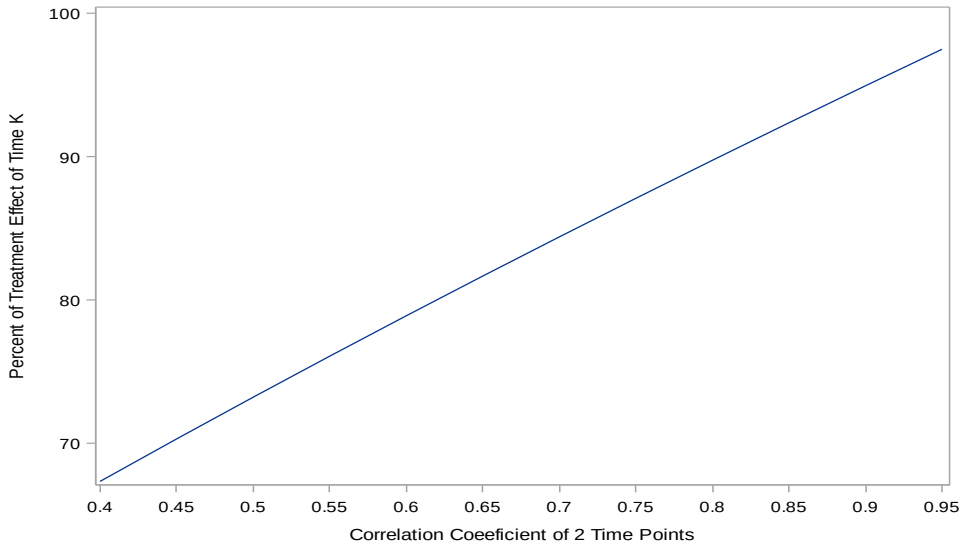
Now we have a similar t-statistics when using the average of time K-1 and K

$$T_2 = \frac{\delta_2}{\sigma_2 \sqrt{\frac{1}{n_a} + \frac{1}{n_p}}}$$

Where treatment effect $\delta_2 = \frac{(p+1)}{2}\delta$ and $\sigma_2 = \sqrt{(1+\rho)/2}\sigma$.

To make $T_2 = T$, we need

$\frac{(p+1)}{2} = \sqrt{(1+\rho)/2}$ or $p = \sqrt{2(1+\rho)} - 1.$ (2)



Line of $T_2 = T$ defined by Percent of δ of Time K and correlation coefficient ρ

For any point (p , ρ) above the curve, $T_2 = T$, the average endpoint has a larger t-statistics than the endpoint defined on the single time point.

For the condition,

$$\frac{(p+1)}{2} = \sqrt{(1+\rho)/2}$$

$\frac{(p+1)}{2}$ indicates the loss of treatment effect due to the average and $\sqrt{(1+\rho)/2}$ indicates the reduced variance due to the average.

For example, if correlation coefficient ρ is 0.6, then p should be greater than 79% of treatment effect of Time K to define an average endpoint of Time K-1 and K.

4. Extend the idea to the endpoint defined on average of more than 2 time points:

For an endpoint defined on average of time K-2, K-1 and K, to get a simple formula, we need to assume that the treatment δ is linearly increasing along the time. For example, if at time point K-2, treatment effect is $p\delta$ and p is assumed 90%. Then treatment effect at time K-2, K-1 and K would be $90\%\delta$, $95\%\delta$ and δ, respectively.

Again we have a similar t-statistics when using the average of time K-2, K-1 and K.

$$T_3 = \frac{\delta_3}{\sigma_3\sqrt{\frac{1}{n_a}+\frac{1}{n_p}}}$$

Where
$\delta_3 = \frac{(p+1)}{2}\delta$ with linearly increasing assumption of treatment effect
$\sigma_3 = \sqrt{(1+2\rho)/3}\sigma$ with assumption that any 2 time points having the variance matrix defined by (1).

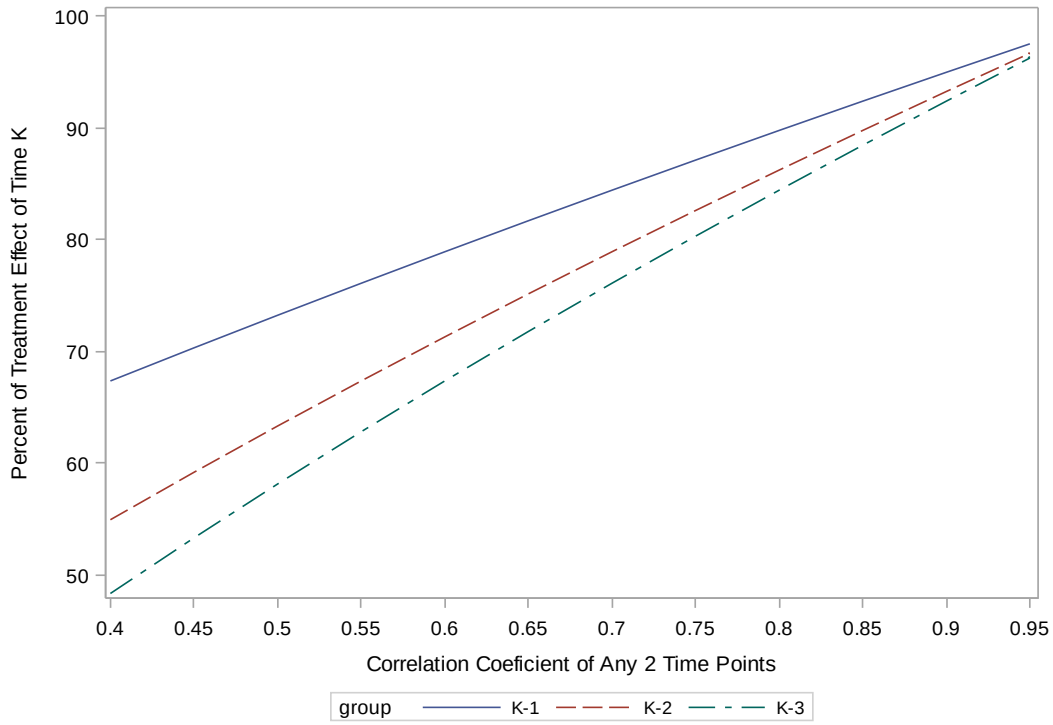
To make $T_3 = T$, we need $p = \sqrt{\frac{4}{3}(1+2\rho)} - 1.$ (3)

For an endpoint defined on average of 4 time points, K-3, K-2, K-1 and K, we have a similar t-statistics .

$$T_4 = \frac{\delta_4}{\sigma_4\sqrt{\frac{1}{n_a}+\frac{1}{n_p}}}$$

Where treatment effect $\delta_4 = \frac{(p+1)}{2}\delta$ with linearly increasing assumption of treatment effect and $\sigma_4 = \sqrt{(1+3\rho)/4}\sigma$ with assumption that any 2 time points having the variance matrix defined by (1).

To make $T_4 = T$, we need $p = \sqrt{(1+3\rho)} - 1.$ (4)



Lines of equal t-statistics defined by Percent of δ of Time K and correlation coefficient ρ

With fixed $\boldsymbol{p}$ and **ρ**, an averages over more time points will have a larger t-statistics.

5. Understanding the Impact of Missing Value

Let us now introduce a parameter **m**, which captures the number of missing values. In this context, we will compare endpoints defined on an average of 2 time points, K-1 and K vs. on single time point K.

Example: Let us assume that n=100 is needed for each group and the drop out rate is 10% leading to a sample size of 110 for each group. If we assume that at time K-1, we have 104 subjects in each group and 100 subjects at time K, then **m** is 4. In short, n is the number of subjects at time K and n+m is the subjects at time K-1. There is no change in the t-statistics if one uses only a single point:

$$\boldsymbol{T} = \frac{\boldsymbol{\delta}}{\boldsymbol{\sigma}\sqrt{\frac{\boldsymbol{1}}{\boldsymbol{n}} + \frac{\boldsymbol{1}}{\boldsymbol{n}}}}$$

But for the endpoint defined on average of K-1 and K, we first partition observations into 2 independent groups:
Group1 include n+n subjects from 2 treatment groups, each has the average of time K-1 and K as the endpoint;
Group2 include m+m subjects from 2 treatment groups, each has only time K-1 as the endpoint;

Now

$$\boldsymbol{T_2 = \frac{\delta_2}{\sqrt{var(\delta_2)}}}$$

Grouping observations into group 1 and 2, we have

$$\delta_2 = \frac{n}{n+m}\frac{(p+1)}{2}\delta + \frac{m}{n+m}p\delta$$

$$var(\delta_2) = (\frac{n}{n+m})^2(\frac{1+\rho}{2})(\frac{1}{n}+\frac{1}{n})\sigma^2 + (\frac{m}{n+m})^2(\frac{1}{m}+\frac{1}{m})\sigma^2 \qquad (5)$$

To make $T_2 = T$, we need (set δ=1, σ=1)

$$\sqrt{\frac{n}{2}} = \frac{\frac{n}{n+m}\frac{(p+1)}{2} + \frac{m}{n+m}p}{\sqrt{(\frac{n}{n+m})^2(\frac{1+\rho}{2})(\frac{2}{n}) + (\frac{m}{n+m})^2(\frac{2}{m})}}$$

$$\sqrt{\frac{n}{2}} = \frac{n\frac{(p+1)}{2} + mp}{\sqrt{(n)^2(\frac{1+\rho}{2})(\frac{2}{n}) + (m)^2(\frac{2}{m})}}$$

$$\sqrt{\frac{n}{2}} = \frac{n(p+1)+2mp}{2\sqrt{n(1+\rho)+2\mathrm{m}}}$$

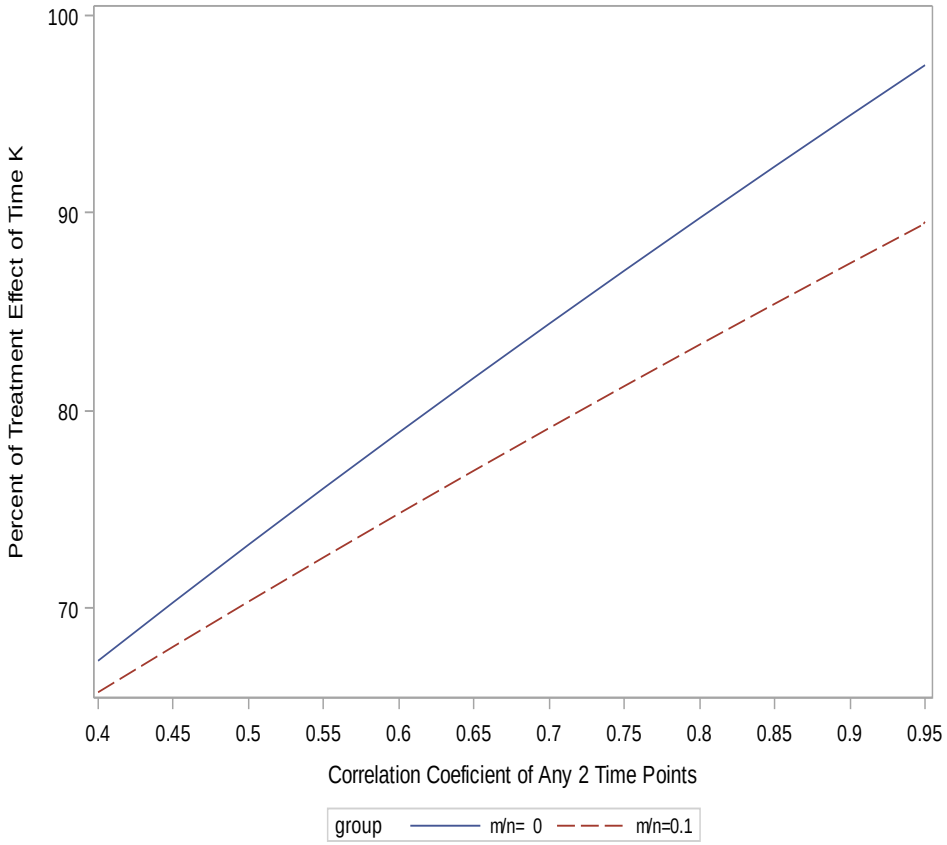
$$2\sqrt{n(1+\rho)+2\mathrm{m}}\sqrt{\frac{n}{2}} = n(p+1)+2mp = p(n+2m)+n$$

$$2\sqrt{n(1+\rho)+2\mathrm{m}}\sqrt{\frac{n}{2}} - n = p(n+2m)$$

$$2\sqrt{(1+\rho)/2+\mathrm{m/n}} - 1 = p(1+2m/n)$$

$$\boldsymbol{p = \frac{\sqrt{2(1+\rho)+4\mathrm{m/n}} - 1}{(1+2m/n)}}$$

If $\mathbf{m = 0}$, then $\boldsymbol{p = \sqrt{2(1+\rho)} - 1}$ as expected.



Line of $T_2 = T$ defined by Percent of δ of Time K and correlation coefficient ρ

6. Summary:

We derived simple formula to select between an endpoint based upon a single time point at the end of trial or upon an average of multiple time points.