Combining Censored and UnCensored Data in a *U*-Statistic: Design and Sample Size Implications for Cell Therapy Research

Lemuel A. Moyé, M.D., Ph.D., University of Texas School of Public Health

Dejian Lai, Ph.D., University of Texas School of Public Health

Kaiyan Jing, University of Texas School of Public Health

Sarah Baraniuk, Ph.D. University of Texas School of Public Health

Minjung Kwak, Ph.D., National Heart, Lung and Blood Institute

Marc S. Penn, M.D., Ph.D., The Cleveland Clinic Foundation

Colin O. Wu, Ph.D, National Heart, Lung and Blood Institute

Cardiovascular Cell Therapy Research Network (CCTRN)

Key words: U-Statistic Clinical Trials Score Function Stem Cells

Abstract

The assumptions that anchor large clinical trials are rooted in smaller, Phase II studies. In addition to specifying the target population, intervention delivery, and patient follow-up duration, physician-scientists who design these Phase II studies must select the appropriate response variables (endpoints). However, endpoint measures can be problematic. If the endpoint assesses the change in a continuous measure over time, then the occurrence of an intervening significant clinical event (SCE) e.g., death can preclude the follow-up measurement. Finally, the ideal continuous endpoint measurement may be contraindicated in a fraction of the study patients, a change that requires a less precise substitution in this subset of participants.

A score function that is based on the *U*-statistic can address these issues of 1) intercurrent SCE's and 2) response variable ascertainments that use different measurements of different precision. The scoring statistic is easy to apply, clinically relevant, and provides flexibility for the investigators' prospective design decisions. Sample size and power formulations for this statistic are provided as functions of clinical event rates and effect size estimates that are easy for investigators to identify and discuss. Examples are provided from current cardiovascular cell therapy research.

1. Introduction

Coronary artery disease (CAD) remains the single largest killer of Americans, producing myocardial infarctions and heart failure (HF).[24] Recent research has delivered substantial improvements in medical therapy and coronary artery revascularization reducing coronary heart disease mortality.[5] However, despite advances in therapy, CAD is a leading cause of HF which bears its own increased morbidity and mortality risks and health costs in an enlarging patient population. Seven million heart attack hospitalizations in the US have generated almost 5 million patients living with HF who face end-stage HF with its 5-year mortality of approximately 50%.[13, 23] Because of the burden faced by these patients with limited options, investigation of alternative treatments are needed. One potential treatment strategy is the use of bone marrow–derived mononuclear cells (BMMNCs), a source of stem cells the treatment of patients with ischemic cardiomyopathy.

The cost and complexity of pivotal clinical trials require that the foundation of these studies be solid. The bedrock of these relatively large experiments is commonly a set of smaller clinical studies that must identify a panoply of possible beneficial and adverse therapy effects through the wise selection of endpoints.

Endpoint selection is challenging in early human cardiovascular cell therapy clinical trials. Possible choices are the size of the heart damaged by a heart attack, known as the infarct region [26, 1] or changes in the percent of blood ejected by the left ventricle with

each heart beat, or left ventricular ejection fraction (EF) [2]. Recent attention has focused on other measures of left ventricular dysfunction e.g., left ventricular end-diastolic volume (LVEDV) (how large the left ventricle becomes at the peak of the cardiac cycle when it is full of blood), and left ventricular end systolic volume (LVESV) (how small the ventricle is when it has ejected its blood content [21]).

Continuous response variables (endpoints) provide necessary statistical power in well designed clinical experiments. However, complications moderate enthusiasm for these continuous endpoints that typically require measurements at both baseline and during the follow-up period. For example, the occurrence of an intervening significant clinical event (SCE) (e.g., death) precludes the follow-up measurement, reducing the precision of the overall measure of therapy effect by reducing the number of evaluable patients. The observation that there may be a greater proportion of patients with an SCE in the control group than in the treatment group introduces an additional informative censoring complication to the analysis. The informative censoring approach of Follmann, Wu, et. al. [4] provides a useful tool for analyzing data in the presence of informative censoring, however, there is no literature on trial design and sample size computations using the informative censoring procedure.

In addition, there can be competing technologies to measure the continuous endpoint. For example a Single Photon Emission Computed Tomography (99TM-SEST SPECT) scan is an imaging test that shows how blood flows to tissues and organs. Using computed tomography (CT) and a radioactive tracer (technetium is the most popular currently), it

reveals how blood flows not just in the chambers of the heart but in and through the heart muscle itself. ^{99}TM-SEST SPECT measures the degree to which damaged heart muscle receives life sustaining blood flow.[9] While some believe that cMR is superior to ^{99}TM-SEST SPECT, cMR measures cannot be obtained in patients who have an implantable (metallic) device e.g., a pace maker. A statistical procedure that permits the more accurate measure of perfusion when it is available, and uses only the less precise measure in its absence allows the use of every patient's data, regardless of the measurement that is indicated by the patient's condition.

This manuscript discusses the development of a *U*-statistic to permit 1) the inclusion of a dichotomous endpoint (SCE) as well as a continuous endpoint in a single primary endpoint, and 2) the use of the more precise endpoint information when it is available, falling back on less precise information when only it can be obtained. Its development and calibration are based on commonly used event rates and measures of efficacy of both the dichotomous and continuous components that clinicians use and understand. Examples are provided from ongoing cardiovascular cell therapy research.

2. Background

The development of new strategies to improve heart function following a heart attack of acute myocardial infarction (AMI) has been a prominent goal for cardiovascular investigation. Studies in animal models have demonstrated that heart function can be significantly improved with bone marrow-derived stem cells following experimental heart at-

tacks induced in animals [20, 10, 7, 34]. Although data supporting significant heart regeneration in these preclinical studies has not been uniform [17, 3], it has led to a number of clinical trials testing the strategy that delivery of of a patient's own (or autologous) bone marrow-derived mononuclear cells (BMMNCs) into the infarct region following AMI may improve heart function [25, 8, 33, 14]. In light of the relative paucity of mechanistic studies into important questions, such as timing of cell delivery, the National Heart, Lung, and Blood Institute (NHLBI) established the Cardiovascular Cell Therapy Research Network (CCTRN) to accelerate research into the use of cell-based therapies for the management of cardiovascular diseases. The Transplantation in Myocardial Infarction Evaluation (TIME) study is a Phase II trial developed by the CCTRN to provide further research into the efficacy, safety, and most appropriate timing of autologous BMMNCs in high-risk, post-AMI patients.

3. Methods

This method is based on two-sample the *U*-statistic [11], a well established, nonparametric measure of effect based on an investigator-determined scoring mechanism. Our development is modeled after the *U*-statistic's implementation to score the occurrence of a combination of two discrete endpoints in a cardiovascular clinical trial [15, 16, 22]. A recent use of this statistic in medical research has been its application to multivariate ordinal data [32].

In its simplest adaptation, the *U*-statistic "builds itself up" from a prospectively selected scoring procedure. Let there be n observations in thecontrol group. Let each of the n patients in the control group have a continuous endpoint measure x_i, $i = 1, 2, 3, \ldots, n$. Similarly, let the primary endpoint measure for each of the kn patients in the active group be indexed by y_j, $j = 1, 2, 3, \ldots, kn$.

The *U*-statistic requires a simple scoring mechanism, denoted by $\varphi_{i,j}$. This is the assignment of a score designed in this paper based on comparing the i^{th} patient in the control group with the j^{th} patient in the active group. The score may be as simple as $\varphi_{i,j} = 1$ if $x_i > y_j$; $\varphi_{i,j} = 0$ if $x_i = y_j$; or $\varphi_{i,j} = -1$ if $x_i < y_j$. Since each of the n control group patients will be compared to each of the kn active group patients, there are kn^2 comparisons. The U-score statistic, W_e is simply the average of these kn^2scores,

$$W_e = \frac{1}{kn^2}\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij} \tag{1}$$

The normalized statistic based on these scores for a test of the null hypothesis (H_0) of no treatment effect versus the alternative hypothesis (H_a) of a change in the distribution of the y_j's based on the treatment is

$$TS = \frac{W_e - \mathrm{E}\left[W_e \mid H_0\right]}{\sqrt{Var\left[W_e \mid H_0\right]}} \tag{2}$$

Under mild regulatory conditions and adequate sample size, we assume that (2) follows a standard normal distribution, then we can compute the sample size from

$$N = 2n = 2\left[\frac{Z_{1-\alpha/2}\sqrt{v_0} \quad - \quad Z_{\beta}\sqrt{v_a}}{\mathrm{E}\left[W_e \mid H_a\right] - \mathrm{E}\left[W_e \mid H_0\right]}\right]^2 \tag{3}$$

where $v_0 = Var\left[W_e \mid H_0\right], v_a = Var\left[W_e \mid H_a\right],$ α is the probability of a type I error, β is the probability of a type II error, and Z_c is the c^{th} percentile value from the standard normal distribution, Alternatively, power may be computed from

$$1-\beta = 1-\Phi_Z\left[\frac{Z_{1-\alpha/2}\sqrt{Var\left(W_e \mid H_a\right)} \quad - \quad \left(\mathrm{E}\left[W_e \mid H_a\right] - \mathrm{E}\left[W_e \mid H_0\right]\right)}{\sqrt{Var\left(W_e \mid H_a\right)}}\right] \tag{4}$$

where $\Phi_Z(z)$ is the cumulative distribution function of the standard normal distribution.

However, the adoption of this statistic requires a careful justification of the scoring mechanism required for the response variables (endpoints). The setting for our evaluations is that of a randomized clinical trial with both a control and an active group. We will construct the score statistic in two cases:

Case 1. A dichotomous right censored measure combined with a single continuous response variable.

Case 2. A dichotomous right censored measure combined with two continuous response variables to be used in a hierarchy determined by the precision of the two response variables.

The mathematics of Case 1 will be developed in detail, and then applied to the Case 2 scenario.

Case 1: A dichotomous right censored measure combined with a single continuous response variable.

The investigators' goal is to compare the change in the measure of a single continuous response variable over time in the control group to the change in that same variable in the active group. This requires that for each patient, there be a measurement at baseline and at the end of the study. However, the investigators recognize that this goal may not be achievable in all patients because of the occurrence of death or another SCE. We will assume that, as is the case with a common response variable in cardiovascular research (e.g., LVEF) an increase in the response variable over time corresponds to improved health status.

Let r be the continuous endpoint variable. Then, for the i^{th} patient in the control group, $i = 1, 2, 3, \ldots, n$, let $d_i(x) = r_{i,2}(x) - r_{i,1}(x)$ be the change in this variable over the duration of the study. Assume that $d_i(x)$ is normally distributed with mean $\mu_{\Delta R}(x)$ and variance σ_x^2. Analogously, let $d_j(y) = r_{j,2}(y) - r_{j,1}(y)$ be the change in the endpoint measure for the j^{th} patient in the active group, which is normally distributed with mean $\mu_{\Delta R}(y)$ and known variance σ_y^2. Under the null hypothesis of the study, $\mu_{\Delta R}(x) = \mu_{\Delta R}(y)$. If we assume that larger values of $\mu_{\Delta R}$ correspond to improved health, then under the alternative hypothesis, the researchers expect that $\mu_{\Delta R}(x) < \mu_{\Delta R}(y)$.

However, the occurrence of a significant event (SCE) (e.g., a death, a recurrent myocardial infarction (MI), can affect the follow-up measurement of the continuous variable. The hallmark of the SCE is that 1) its occurrence during the trial either precludes the follow-up measurement (as in the case of death), or perturbs the measurement to the point that the effect of therapy can be difficult to assess (e.g., the occurrence of an intercurrent heart attack), and 2) the SCE event rates in the randomized groups may themselves be related to the therapy effect. The occurrence of an intervening SCE (itself an underpowered evaluation in a small study) reduces the power of the LVEF measure by decreasing the number of patients who survive to have the follow-up measurement.

In this case we define the scoring mechanism $\varphi_{i,j}$ as follows:

$\varphi_{i,j} = 1$ if both the i^{th} patient if the control group and the j^{th} patient in the active group experience an SCE during the study, and the time to event for the control group patient is less than the time to event for the active group patient.

$\varphi_{i,j} = 1$ if the i^{th} patient in the control group experiences an SCE during the study and the j^{th} patient in the active group does not experience an SCE during the study.

$\varphi_{i,j} = -1$ if both the i^{th} patient in the control group and the j^{th} patient in the active group experiences an SCE during the course of the study, but the time to event for the control group patient is greater than the time to event for the active group patient.

$\varphi_{i,j} = -1$ if the i^{th} patient in the control group does not experience an SCE during the study and the j^{th} patient in the active group does experience an SCE during the study.

$\varphi_{i,j} = c$ if neither the i^{th} patient in the control group nor the j^{th} patient in the active group experience an SCE during the study, and the change in the continuous measure r for the control group patient is less than the change in continuous measure for the active group patient.

$\varphi_{i,j} = -c$ if neither the i^{th} patient in the control group nor the j^{th} patient in the active group experiences an SCE during the study, and the change in the continuous measure r_i for the control group patient is greater than the change in continuous measure for the active group patient, r_j.

$\varphi_{i,j} = 0$ otherwise.

Under this mechanism, the occurrence of an early SCE (e.g., a death) in one group is considered worse than a patient survival or a later occurring SCE in the other treatment group. If both patients in the comparison have no SCE, then the change in the response variable is compared.

With some additional notation, the assignment of this scoring system permits the computation of the mean and variance of W_e under the null and alternative hypothesis.

Notation:

Define $C_{X(i)}(E, R)$ as the endpoint status of the i^{th} patient in the control group, and $C_{Y(j)}(E, R)$ as the endpoint status of the j^{th} patient in the active group. We will use this notation to allow us to capture either 1) the time to the occurrence of an SCE if one has occurred dur-

ing the course of the trial, or 2) the change in the continuous variable if an SCE has not occurred.

If an SCE has occurred for the i^{th} patient in the control group, then $C_{X(i)}(E, R) = C_{X(i)}(+, R)$, and its value is the time to the occurrence of the SCE. Since the SCE has occurred during the course of the study, then $0 \le C_{X(i)}(+, R) \le T$ where T is the maximum time a patient is to be followed in the research protocol. If an SCE has not occurred, then $C_{X(i)}(E, R) = C_{X(i)}(-, R)$, and we set $C_{X(i)}(-, R)$ to equal the change in the continuous measure. Identical notation applies to the j^{th} patient in the active group, $C_{Y(j)}(E, R)$.

For example, if in a 180 day clinical trial, the 4th patient in the control group died on day 117, then $C_{X(4)}(E, R) = C_{X(4)}(+, R) = 117$, the positive sign signifying that the SCE event occurred. Alternatively, if the 5th patient in the active group survived the trial and experienced a six unit increase in the continuous response variable, then $C_{Y(5)}(E, R) = C_{Y(5)}(-, R) = 6$, the minus sign in $C_{Y(5)}(-, R)$ indicating that no SCE occurred during the study.

Using this notation, and letting $\mathbf{1}_{X \in A}$ be the indicator function that takes the value of 1 when x is a member of set A and 0 otherwise, we can write the score function $\varphi_{i,j}$ as

$$\begin{aligned}\varphi_{i,j} = {} & \mathbf{1}_{\left[C_{x(i)}[+,R] < C_{y(j)}[+,R]\right]} + \mathbf{1}_{\left[C_{x(i)}[E,R]=C_{x(i)}[+,R] \cap C_{x(i)}[E,R]=C_{y(j)}[-,R]\right]} - \mathbf{1}_{\left[C_{x(i)}[+,R] > C_{y(j)}[+,R]\right]} \\ & - \mathbf{1}_{\left[C_{x(i)}[E,R]=C_{x(i)}[-,R] \cap C_{x(i)}[E,R]=C_{y(j)}[+,R]\right]} + c\mathbf{1}_{\left[C_{x(i)}[-,R] < C_{y(j)}[-,R]\right]} - c\mathbf{1}_{\left[C_{x(i)}[-,R] > C_{y(j)}[-,R]\right]}.\end{aligned} \tag{5}$$

3.1 Computing $E\left[\varphi_{i,j}\right]$

The notation from the previous section permits us to write the expected value of $\varphi_{i,j}$ under the hypothesis $H_{k,}$ $k = 0$ for the null hypothesis, and $k = a$ for the alternative hypothesis. We assume throughout this manuscript that the time to an SCE and the continuous measure are independent.

$$\begin{aligned}\boldsymbol{E}\left[\varphi_{ij} \mid H_k\right] &= \boldsymbol{P}\left[C_{x(i)}[+,R] < C_{y(j)}[+,R] \mid H_k\right] \\ &+ \boldsymbol{P}\left[C_{x(i)}[E,R] = C_{x(i)}[+,R] \cap C_{x(i)}[E,R] = C_{y(j)}[-,R] \mid H_k\right] \\ &- \boldsymbol{P}\left[C_{x(i)}[+,R] > C_{y(j)}[+,R] \mid H_k\right] - \boldsymbol{P}\left[C_{x(i)}[E,R] = C_{x(i)}[-,R] \cap C_{x(i)}[E,R] = C_{y(j)}[+,R] \mid H_k\right] \\ &+ c\boldsymbol{P}\left[C_{x(i)}[-,R] < C_{y(j)}[-,R] \mid H_k\right] - c\boldsymbol{P}\left[C_{x(i)}[-,R] > C_{y(j)}[-,R] \mid H_k\right].\end{aligned} \tag{6}$$

This computation is straightforward when the probability distributions of 1) the occurrence of SCE's and 2) the probability distribution of the continuous response variable *r* are known. For example, assume the time to an SCE follows an exponential distribution with parameter λ_x in the control group and λ_y in the active group. Also assume that the change in the continuous measure *r* follows a normal distribution with mean as before $\mu_{\Delta R}(x)$ and standard deviation $\sigma_{\Delta R}(x)$ in the control group, and analogously mean $\mu_{\Delta R}(y)$, and standard deviation $\sigma_{\Delta R}(y)$ in the active group. The first term on the right hand side of equation (6) is

$$\mathbf{P}\left[C_{x(i)}[+,R] < C_{Y(j)}[+,R]\right] = \int_0^T \lambda_y e^{-\lambda_y y} \int_0^y \lambda_x e^{-\lambda_x x} dxdy = 1 - e^{-\lambda_y T} - \frac{\lambda_y}{\lambda_x + \lambda_y}\left(1 - e^{-(\lambda_x + \lambda_y)T}\right)$$

As another example, the last term on the right hand side of equation (6) is

$$\mathbf{P}\left[C_{x(i)}[-,R] < C_{Y(j)}[-,R] \mid H_0\right] = e^{-(\lambda_x + \lambda_y)T} \Phi_Z\left[\frac{-\left(\mu_{\Delta R}(x) - \mu_{\Delta R}(y)\right)}{\sqrt{\sigma^2_{\Delta R}(x) + \sigma^2_{\Delta R}(y)}}\right] \tag{7}$$

Since the null hypothesis assumes no treatment effect, we let $\lambda_x = \lambda_y$, and $\mu_{\Delta R}(x) = \mu_{\Delta R}(y)$, to see that

$$\mathbf{E}\left[\varphi_{ij} \mid H_0\right] = 0. \tag{8}$$

Under the alternative hypothesis H_a of a treatment effect, then either $\lambda_x \neq \lambda_y$, and/or $\mu_{\Delta R}(x) \neq \mu_{\Delta R}(y)$, permitting us to write,

$$\begin{aligned}\mathbf{E}\left[\varphi_{ij} \mid H_a\right] = {} & 1 - e^{-\lambda_y T} - \frac{\lambda_y}{\lambda_x + \lambda_y}\left(1 - e^{-(\lambda_x+\lambda_y)T}\right) + e^{-\lambda_y T}\left(1 - e^{-\lambda_x T}\right) \\ & - \left[1 - e^{-\lambda_x T} - \frac{\lambda_x}{\lambda_x + \lambda_y}\left(1 - e^{-(\lambda_x+\lambda_y)T}\right) + e^{-\lambda_x T}\left(1 - e^{-\lambda_y T}\right)\right] \\ & + ce^{-(\lambda_x+\lambda_y)T}\Phi_Z\left[\frac{-\left(\mu_{\Delta R}(x) - \mu_{\Delta R}(y)\right)}{\sqrt{\sigma^2_{\Delta R_X} + \sigma^2_{\Delta R_Y}}}\right] - ce^{-(\lambda_x+\lambda_y)T}\left(1 - \Phi_Z\left[\frac{-\left(\mu_{\Delta R}(x) - \mu_{\Delta R}(y)\right)}{\sqrt{\sigma^2_{\Delta R_X} + \sigma^2_{\Delta R_Y}}}\right]\right)\end{aligned} \tag{9}$$

3.2 Variance computation

Computation of the variance of W_e, while somewhat more complicated than the mean, is executable. Assume n subjects in the control group and kn subjects in the active group, then,

$$\mathbf{Var}\left[W_e\right] = \mathbf{Var}\left[\frac{1}{kn^2}\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right] = \frac{1}{k^2 n^4}\mathbf{Var}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right].$$

Further,

$$\mathbf{Var}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right]=\mathbf{E}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right)^2\right]-\left(\mathbf{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right]\right)^2. \tag{10}$$

The last term on the right of the second line of (10) is easily evaluated.

$$\left(\mathbf{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right]\right)^2=\left(kn^2\mathbf{E}\left[\varphi_{ij}\right]\right)^2=k^2n^4\mathbf{E}^2\left[\varphi_{ij}\right].$$

where the expected value of $\mathbf{E}\left[\varphi_{ij}\right]$ has already been computed both under the null (8) and alternative **(9)** hypotheses.

To evaluate $\mathbf{E}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right)^2\right]$ from (10), we rewrite $\left(\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right)^2$ as

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right)^2=\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq j'}}^{kn}\varphi_{ij}\varphi_{ij'}+\sum_{\substack{i=1\\ i\neq i'}}^{n}\sum_{j=1}^{kn}\varphi_{ij}\varphi_{i'j}+\sum_{\substack{i=1\\ i\neq i'}}^{n}\sum_{\substack{j=1\\ j\neq j'}}^{kn}\varphi_{ij}\varphi_{i'j'} \tag{11}$$

This helpful simplification is due to Gehan [6]. We may now pass the expectation argument through the preceding equation to find

$$\begin{aligned}\mathbf{E}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}\right)^2\right]&=\mathbf{E}\left[\sum_{i=1}^{kn}\sum_{j=1}^{n}\varphi_{ij}^2\right]+\mathbf{E}\left[\sum_{i=1}^{kn}\sum_{\substack{i'=1\\ i\neq i'}}^{kn}\sum_{j=1}^{n}\varphi_{ij}\varphi_{i'j}\right]\\&+\mathbf{E}\left[\sum_{i=1}^{kn}\sum_{j=1}^{n}\sum_{\substack{j'=1\\ j\neq j'}}^{n}\varphi_{ij}\varphi_{ij'}\right]+\mathbf{E}\left[\sum_{i=1}^{kn}\sum_{\substack{i'=1\\ i\neq i'}}^{kn}\sum_{j=1}^{n}\sum_{\substack{j'=1\\ j\neq j'}}^{n}\varphi_{ij}\varphi_{i'j'}\right]\end{aligned} \tag{12}$$

and evaluating term by term we see

$$\mathbf{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}^{2}\right]=kn^{2}\mathbf{E}\left[\varphi_{ij}^{2}\right]$$

$$\mathbf{E}\left[\sum_{i=1}^{n}\sum_{\substack{i'=1\\ i'\neq i}}^{n}\sum_{j=1}^{kn}\varphi_{ij}\varphi_{i'j}\right]=kn^{2}\left(n-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]$$

$$\mathbf{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\sum_{\substack{j=1\\ j\neq j'}}^{kn}\varphi_{ij}\varphi_{ij'}\right]=kn^{2}\left(kn-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$$

$$\mathbf{E}\left[\sum_{i=1}^{n}\sum_{\substack{i'=1\\ i'\neq i}}^{n}\sum_{j=1}^{kn}\sum_{\substack{j=1\\ j\neq j'}}^{kn}\varphi_{ij}\varphi_{i'j'}\right]=kn^{2}\left(kn-1\right)\left(n-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{i'j'}\right]$$

$\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$ is the expected value of the product of the scoring function between 1) the i^{th} control group patient and the j^{th} active group patient, and 2) the same i^{th} control group patient but a different j'^{th} active group patient where $j\neq j'$. $\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]$ is an analogous computation involving the i^{th} and i'^{th} in the control group and the j^{th} patient in the active group.We may now rewrite (12) as

$$\mathbf{E}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi_{ij}\right)^{2}\right]=kn^{2}\mathbf{E}\left[\varphi_{ij}^{2}\right]+kn^{2}\left(n-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]$$
$$+kn^{2}\left(kn-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]+kn^{2}\left(kn-1\right)\left(n-1\right)^{2}\mathbf{E}^{2}\left[\varphi_{ij}\right]$$

Thus

$$\begin{aligned}&\mathbf{Var}\left[\sum_{i=1}^{kn}\sum_{j=1}^{n}\varphi_{ij}\right]\\&=kn^{2}\mathbf{E}\left[\varphi_{ij}^{2}\right]+kn^{2}\left(n-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]+kn^{2}\left(kn-1\right)\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]\\&+kn^{2}\left(kn-1\right)\left(n-1\right)^{2}\mathbf{E}^{2}\left[\varphi_{ij}\right]-k^{2}n^{4}\mathbf{E}^{2}\left[\varphi_{ij}\right]\end{aligned}\qquad(13)$$

Further simplication reveals

$$
\begin{aligned}
\mathbf{Var}\left[W_e\right] &= \frac{1}{k^2 n^4}\mathbf{Var}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right] \\
&= \frac{n^3}{k^2 n^4}\; k\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right] + k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right] - \left(k+1\right)\mathbf{E}^2\left[\varphi_{ij}\right]\right) \\
&- \frac{n^2}{k^2 n^4}\left(k\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right] + \mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right] - \mathbf{E}\left[\varphi_{ij}^2\right] - \mathbf{E}^2\left[\varphi_{ij}\right]\right)\right) \\
&= \frac{1}{k^2}\left[\frac{A}{n} \;-\; \frac{B}{n^2}\right]
\end{aligned}
\tag{14}
$$

Since the variance can be computed under both the null and alternative hypothesis, we may write $\mathbf{Var}\left[W_e|H_0\right] = V_0 = \frac{1}{k^2}\left[\frac{A_0}{n} \;-\; \frac{B_0}{n^2}\right]$ and $\mathbf{Var}\left[W_e|H_a\right] = V_a = \frac{1}{k^2}\left[\frac{A_a}{n} \;-\; \frac{B_a}{n^2}\right]$

where A_0 and B_0 are computed under the null hypothesis and A_a and B_a are computed under the alternative.

3.3 Sample size computation

Assuming a normal distribution for W_e, we compute that,

$$
TS = \frac{W_e - \mathbf{E}\left[W_e \mid H_0\right]}{\sqrt{\mathbf{Var}\left[W_e|H_0\right]}} \tag{15}
$$

And from consideration of the type II error, we may write,

$$
Z_\beta = \frac{Z_{1-\alpha/2}\sqrt{\mathbf{Var}\left[W_e|H_0\right]} - \mathbf{E}\left[W_e \mid H_a\right]}{\sqrt{\mathbf{Var}\left[W_e|H_a\right]}} \tag{16}
$$

and

$$
Z_\beta\sqrt{\mathbf{Var}\left[W_e|H_a\right]} = Z_{1-\alpha/2}\sqrt{\mathbf{Var}\left[W_e|H_0\right]} - \mathbf{E}\left[W_e \mid H_a\right]. \tag{17}
$$

Substituting for the variance term, we may write

$$Z_{1-\alpha/2}\sqrt{\frac{1}{k^2}\left[\frac{A_0}{n} - \frac{B_0}{n^2}\right]} - Z_{\beta}\sqrt{\frac{1}{k^2}\left[\frac{A_a}{n} - \frac{B_a}{n^2}\right]} = \mathbf{E}\left[W_e \mid H_a\right]. \tag{18}$$

Squaring both sides, simplifying, and squaring again, with expansion and further simplication produces the quartic equation

$$a_4 n^4 + a_3 n^3 + a_2 n^2 + a_1 n + a_0 = 0 \tag{19}$$

where

$$\begin{aligned}
a_4 &= k^4 \mathbf{E}^4\left[W_e \mid H_a\right] \\
a_3 &= -2\left(Z^2_{1-\alpha/2}\mathbf{E}^2\left[W_e \mid H_a\right]k^2 A_0 + Z^2_{\beta}\mathbf{E}^2\left[W_e \mid H_a\right]k^2 A_a\right) \\
a_2 &= Z^4_{1-\alpha/2}A_0^2 + Z^4_{\beta}A_a^2 + 2Z^2_{1-\alpha/2}Z^2_{\beta}A_0^2A_a^2 + 2Z^2_{1-\alpha/2}\mathbf{E}^2\left[W_e \mid H_a\right]k^2 B_0 \\
&\quad + 2Z^2_{\beta}\mathbf{E}^2\left[W_e \mid H_a\right]k^2 B_a - 4Z^2_{1-\alpha/2}Z^2_{\beta}A_0^2A_a^2 \\
a_1 &= -2Z^4_{1-\alpha/2}A_0B_0 - 2Z^4_{\beta}A_aB_a - 2Z^2_{1-\alpha/2}Z^2_{\beta}A_0B_a - 2Z^2_{1-\alpha/2}Z^2_{\beta}A_aB_0 \\
&\quad + 4Z^2_{1-\alpha/2}Z^2_{\beta}A_0B_a + 4Z^2_{1-\alpha/2}Z^2_{\beta}A_aB_0 \\
a_0 &= Z^4_{1-\alpha/2}B_0^2 + Z^4_{\beta}B_a^2 + 2Z^2_{1-\alpha/2}Z^2_{\beta}B_0B_a - 4Z^2_{1-\alpha/2}Z^2_{\beta}B_0B_a
\end{aligned} \tag{20}$$

Power can be more directly computed as

$$1 - \Phi_Z\left[\frac{Z_{1-\alpha/2}\sqrt{\frac{1}{k^2}\left[\frac{A_0}{n} - \frac{B_0}{n^2}\right]} - \mathbf{E}\left[W_e \mid H_a\right]}{\sqrt{\frac{1}{k^2}\left[\frac{A_a}{n} - \frac{B_a}{n^2}\right]}}\right] \tag{21}$$

3.4 Asymptotic approach for $\mathbf{Var}\left[W_e\right]$:

Working from equation (14)

$$\begin{aligned}
\mathbf{Var}\left[W_e\right] &= \frac{1}{k^2n^4}\mathbf{Var}\left[\sum_{i=1}^{n}\sum_{j=1}^{kn}\varphi_{ij}\right] \\
&= \frac{n^3}{kn^4}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right] + k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right] - (k+1)\mathbf{E}^2\left[\varphi_{ij}\right]\right) \\
&\quad - \frac{n^2}{kn^4}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right] + \mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right] - \mathbf{E}\left[\varphi_{ij}^2\right] - \mathbf{E}^2\left[\varphi_{ij}\right]\right)
\end{aligned}$$

Ignoring terms on the order of n^{-2}, we have

$$\mathbf{Var}\left[W_e\right]=\frac{1}{kn}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]+k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]-(k+1)\mathbf{E}^2\left[\varphi_{ij}\right]\right), \tag{22}$$

Equation (22) is the variance of W_e under the alternative hypothesis, $\mathbf{Var}\left[W_e \mid H_a\right]$. Under the null hypothesis, we assume $\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]=\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$, and $\mathbf{E}\left[\varphi_{ij}\right]=0$.

$\mathbf{Var}\left[W_e \mid H_0\right]=\frac{k+1}{kn}\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right]$. We may write

$$\begin{aligned}\mathbf{Var}\left[W_e \mid H_0\right]&=\frac{k+1}{kn}\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right]\\ \mathbf{Var}\left[W_e \mid H_a\right]&=\frac{1}{kn}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]+k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]-(k+1)\mathbf{E}^2\left[\varphi_{ij}\right]\right)\end{aligned} \tag{23}$$

Now substituting equations for the $\mathbf{Var}\left[W_e \mid H_0\right]$ and $\mathbf{Var}\left[W_e \mid H_a\right]$ from (23) into (17) to compute the sample size of the trial, we write

$$\begin{aligned}&Z_\beta\sqrt{\frac{1}{kn}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_a\right]+k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_a\right]-(k+1)\mathbf{E}^2\left[\varphi_{ij} \mid H_a\right]\right)}\\ &=Z_{1-\alpha/2}\sqrt{\frac{k+1}{kn}\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right]}-\mathbf{E}\left[W_e \mid H_a\right]\end{aligned} \tag{24}$$

Noting that the total number patients in the study is n control group plus kn in the active group, we can write

$$N=(k+1)\left[\frac{\left(Z_{1-\alpha/2}\sqrt{\frac{k+1}{kn}\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right]}-Z_\beta\sqrt{\frac{1}{kn}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_a\right]+k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_a\right]-(k+1)\mathbf{E}^2\left[\varphi_{ij} \mid H_a\right]\right)}\right)}{\mathbf{E}\left[W_e \mid H_a\right]}\right]^2 \tag{25}$$

Power can be expressed as

$$1-\beta = 1-\Phi_Z\left[\frac{Z_{1-\alpha/2}\sqrt{\frac{k+1}{kn}\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\mid H_0\right]} \;-\; \mathbf{E}\left[W_e \mid H_a\right]}{\sqrt{\frac{1}{k}\left(\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\mid H_a\right] + k\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\mid H_a\right] - (k+1)\mathbf{E}^2\left[\varphi_{ij}\mid H_a\right]\right)}}\right] \tag{26}$$

The structure for $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$ can be identified and tabulated (Table 1) revealing 18 terms, each of which is evaluated under the null and alternative hypothesis. For example one of the terms may be written as,

$$\mathbf{P}\left[C_{X(i)}(+,R) < C_{Y(j)}(+,R) \cap C_{Y(j')}(E,R) = C_{Y(j')}(-,R)\right] \;=\; \mathbf{P}\left[U < V < T < W\right]. \tag{27}$$

Here U follows an exponential distribution with parameter λ_x, T is the duration of the study, and V and W are i.i.d. exponentially distributed random variables with parameter λ_y.. Note that the final expressions for the expectation are in terms of the parameters λ_x, λ_y, $\mu_{\Delta R}(x)$, $\mu_{\Delta R}(y)$, $\sigma^2_{\Delta R}(x)$, and $\sigma^2_{\Delta R}(y)$. These are available from the clinical scientists. Thus (27) may be written as

$$\begin{aligned}&\mathbf{P}\left[C_{X(i)}(+,R) < C_{Y(j)}(+,R) \cap C_{Y(j')}(E,R) = C_{Y(j')}(-,R)\right] \\ &= e^{-\lambda_y T}\left[\left(1-e^{-\lambda_y T}\right) - \frac{\lambda_y}{\lambda_x+\lambda_y}\left(1-e^{-(\lambda_x+\lambda_y)T}\right)\right].\end{aligned} \tag{28}$$

Which may be evaluated under the null hypothesis where $\lambda_x = \lambda_y$, or the alternative where $\lambda_x \neq \lambda_y$.

However, terms that involve comparison of the continuous response variable between three patients must be handled differently. Consider the circumstance where the i^{th} patient in the control group's LVEF has increased by more than the j^{th} and the j'$^{\text{th}}$ patient in the active group. Then one of the expressions required for $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$ is

$$\begin{aligned}&\mathbf{P}\left[C_{X(i)}(-,R) < C_{Y(j)}(-,R) \cap C_{X(i)}(-,R) < C_{Y(j')}(-,R)\right] \\ &= e^{-(\lambda_x+2\lambda_y)T}\mathbf{P}\left[\left(d_i < d_j\right) \cap \left(d_i < d_{j'}\right)\right].\end{aligned} \tag{29}$$

where d_i is the change in the response variable for the i^{th} patient in the control group over the duration of the study, and d_j and $d_{j'}$ are the response variable changes for the j^{th} and j'^{th} patients in the active group respectively. The expression $e^{-(\lambda_x+2\lambda_y)T}$ is the probability that all three patients (one in the control group and two in the active group) have no SCE throughout the course of the trial and therefore have the continuous measure assessed at baseline and at the end of the study.

To compute $\mathbf{P}\left[\left(d_i < d_j\right) \cap \left(d_i < d_{j'}\right)\right]$, recall that for the one control group patient, $d_i \sim N\left(\mu_{\Delta R}(x), \sigma^2_{\Delta R}(x)\right)$, and for the two active group patients, d_j, and $d_{j'}$ are identically distributed as $N\left(\mu_{\Delta R}(y), \sigma^2_{\Delta R}(y)\right)$. If we define the random variables U and V in the affine transformation

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} d_i \\ d_j \\ d_{j'} \end{pmatrix} = \begin{pmatrix} d_i - d_j \\ d_i - d_{j'} \end{pmatrix} \tag{30}$$

We may then write, $\mathbf{P}\left[\left(d_i < d_j\right) \cap \left(d_i < d_{j'}\right)\right] = \mathbf{P}\left[U < 0 \ \cap \ V < 0\right]$. Since the joint distribution of $r_i(x)$, $r_j(y)$, and $r_{j'}(y)$, is multivariate normal with mean vector $\underline{\mathbf{u}}$ and variance $\boldsymbol{\Sigma}$,

$$\underline{\boldsymbol{\mu}} = \begin{pmatrix} \mu_{\Delta R}(x) \\ \mu_{\Delta R}(y) \\ \mu_{\Delta R}(y) \end{pmatrix}; \qquad \Sigma = \begin{pmatrix} \sigma^2_{\Delta R}(x) & 0 & 0 \\ 0 & \sigma^2_{\Delta R}(y) & 0 \\ 0 & 0 & \sigma^2_{\Delta R}(y) \end{pmatrix} \tag{31}$$

Then using the transformation of (30), we see that the joint distribution of U and V is bivariate normal with mean and variance

$$\begin{pmatrix} U \\ V \end{pmatrix} \sim \mathbf{MVN}_2\left(\begin{pmatrix} \mu_{\Delta R}(x) - \mu_{\Delta R}(y) \\ \mu_{\Delta R}(x) - \mu_{\Delta R}(y) \end{pmatrix}, \begin{pmatrix} \sigma^2_{\Delta R}(x) + \sigma^2_{\Delta R}(y) & \sigma^2_{\Delta R}(x) \\ \sigma^2_{\Delta R}(x) & \sigma^2_{\Delta R}(x) + \sigma^2_{\Delta R}(y) \end{pmatrix} \right) \tag{32}$$

Thus, the desired probability $\mathbf{P}\left[U < 0 \cap V < 0\right]$ is simply the evaluation of this region over the bivariate distribution defined in (32), and the probability required by (29) is therefore available.

Each of the eighteen terms is computed similarly, and assembled in accordance with Table 1 to construct $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$ under the null and alternative hypotheses. An analogous table can be constructed for computing $\mathbf{E}\left[\varphi_{ij}\varphi_{i'j}\right]$. Then $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right]$ and $\mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_0\right]$ can be substituted into equation (23) to compute the $\mathbf{Var}\left[W_e \mid H_0\right]$ and analogously, $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_a\right]$ and $\mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_a\right]$ will be used to compute $\mathbf{Var}\left[W_e \mid H_a\right]$. Alternatively, $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_0\right], \mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_0\right]$ and $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'} \mid H_a\right], \mathbf{E}\left[\varphi_{ij}\varphi_{i'j} \mid H_a\right]$ can be substituted into equation (20) and (19) to compute the sample size for the exact computation, or equation (25) in the case of asymptotic solution. These two expectations can be used to compute the power in equations (21) and (26) for the exact and asymptotic solutions respectively.

Case 2. Competing clinical measures of different precision

This second case adds one level of complexity to Case 1. The investigators' goal is to compare the change in a continuous measure over time in the control group to that of the active group. However in this circumstance the investigators have two competing assessments of the same endpoint continuous variable. The first, denoted by the continuous variable *r*, is the most precise but is not available in all subjects. The second, denoted by *s*, is less accurate, but is available for everyone.

For this case, we can modify the scoring function from Case 1 so that

$\varphi_{ij} = d$ if neither the i^{th} patient in the control group nor the j^{th} patient in the active group experience an SCE during the study, the change in the variable *r* is not available for both control and active group patients and the change in *s* in control group patient is less than the change in *s* in the active group patient.

$\varphi_{ij} = -d$ if neither the i^{th} patient in the control group nor the j^{th} patient in the active group experience an SCE during the study, the change in the variable *r* is not available for both active and control group patients and the change in *s* in control group patient is greater than the change in *s* in the active group patient.

The computations follow the development of Case 1. The structure for $\mathbf{E}\left[\varphi_{ij}\varphi_{ij'}\right]$, now tabulated requires 22 terms.

4. Results

A series of evaluations of this *U*-statistic in when clinical measures of event rates and the effect of therapy on the continuous variable were carried out (Figures 1 and 2).

Figure 1 and Figure 2 appear here

Figure 1 identifies the relationship between the trial size (total number of patients in both the active and the placebo group) and the probability of a significant clinical event as a function of the effect of cell therapy on the significant clinical event rate as a function of *c*, the weight ascribed to the continuous endpoint measure in the analysis. In this circumstance we assume that the change in the response variable in the active group is five units greater than the change in the control group. We also assume the standard deviation of this change is 7 for each group (80% power and a type I error rate a two sided alpha of 0.05 is assumed for all analyses). In each of the curves in Figure 1, curves, the trial size is larger for larger probabilities of an SCE. Larger probabilities of an SCE increase the proportion of patients who have no measure of the continuous endpoint that is obtained at the conclusion of the study, and larger sample sizes are required in order to main the power of the evaluation of the therapy's impact on the continuous measure.

We also note the sample size increases as the value of c decreases. The value of c is the relative weight in the scoring system. As *c* decreases the impact of a nonzero comparison between the active and control group measures has less weight than that of the comparison of SCE timings. This diminished weight for comparison generates the need for more continuous measure comparisons in the cohort, thereby increasing the sample

size. Figure 2 demonstrates the same effect of decreasing sample size for larger values of the continuous weighting function *c* as a function of the efficacy of the SCE rate. Note that for all values of efficacy evaluated, the sample size stabilized for values of *c* greater than 3.

The well established relationship between sample size and treatment effect (Δ) are demonstrated in Figure 3. In the paradigm of combining a continuous and a dichotomous endpoint, the sample size decreases as the effect size increases, and increases as the treatment standard deviation of the difference (σ_Δ) increases. Analogously, it is well accepted that when a dichotomous measure is used as a response varible in a clinical trial, the trial size increases as the prevalence of the dichotomous variable increases and decreases with increasing efficacy of treatment against that response variable. This is demonstrated in Figure 4.

Figure 3 and Figure 4 appears here

However, the *U*-statistic also includes the impact of the efficacy of therapy on the SCE rate that is built in to the score statistic. In Figure 1, *e* or efficacy is a measure of the percent decrease in the SCE control group event rate p_c generated by the therapy and observed in the active group p_t. The larger the efficacy, the greater the impact of the treatment group on the score statistic. Thus, the score statistic is a function of the effect of the cell therapy on the continuous measure, as reflected by Δ and σ(Δ), and also by the efficacy of the of the therapy on the SCE rate as well *e*. Thus, while larger values of the

probability of an SCE still produce larger trial sizes, the efficacy of the therapy on the SCE rate moderates this relationship.

Case 2.

In this research scenario there are two continuous measures, each with weights *c* and *d*. As both *c* and *d* increase, the weight of each continuous endpoint increases, and the sample size decreases. However, the larger values of *c* and *d* have diminishing impact on the sample size.

(Figure 5 approximately here)

5. Discussion

This manuscript demonstrates a method to combine prospectively declared mortality measures with continuous endpoints that maintains the clinical hierarchy of the occurrence of events, using information from the continuous effect size. No imputation is required, and the difficulties with worse rank assignments to missing continuous endpoint data are avoided.

This circumstance is distinct from the multiple endpoint scenario, where investigators choose from among several different endpoint measures. Many important contributions to the literature have addressed this complex challenge. The multiple testing dilemma has been central to clinical trial interpretation. Clinical trialists commonly face the issue of

endpoint selection and cannot resolve it in the favor of one or the other. Clinical trials can have endpoints with no priority among their selection at all [31, 18]. O'Brien examined the role of a rank sum test in 34 endpoint setting [19]. Lachin suggested the use of imputation, assigning a worst rank score to those patients who are missing the continuous endpoint measure due to a mortal event [12]. Other authors have proposed alternative solutions [19, 27, 28, 29, 30].

Of particular use are the weighting values c and d. The investigator has complete control over the values of these weights but must choose them carefully. For example, in clinical trials in which the predominant response value is dichotomous a weighting score e.g., $0 \leq d \leq c \leq 1$ is attractive. Since the dichotomous variable occurs so frequently (e.g., mortality) and is only replaced by the continuous measures in the cases where vital status information is not available, discounting the contribution of the continuous variables is appropriate. However, in studies, such as smaller cell therapy studies where the response variable is continuous, and relatively small numbers of subjects have SCE's, a greater weight for the continuous measure can be justified. In our cell therapy studies, the value of $c = 4$ is appropriate. We advocate selecting d such that $0 \leq d < c$ since the less precise measure should have less influence on the test statistic than the more precise one. However, these values must be chosen before any endpoint analysis takes place to avoid selections that are biased by the investigators observations of the values of the final response variables.

Three clinical trials in the NHLBI sponsored Cardiovascular Cell Therapy Research Network (CCTRN) are currently underway in which we will assess the utility of this approach. Complications of the application of this procedure include the observation that the event rates of the significant clinical event and the standard deviation of the continuous measure differs from that assumed during the study's design phase. In addition, interim review of the statistics by Data Safety and Monitoring Boards introduces new complications. Neither of these are assessed specifically in this manuscript.

References

1. Abdel-Latif A, Bolli R, Tleyjeh IM, Montori VM, Perin EC, Hornung CA, Zuba-Surma EK, Al-Mallah M,Dawn B. Adult bone marrow-derived cells for cardiac repair. Arch Intern Med. (2007). 167:989-997.

2. Assmus B, Schachinger V, Teupe C, et al. Transplantation of progenitor cells and regeneration enhancement in acute myocardial infarction (TOPCARE-AMI). Circulation 2002;106:3009-3017.

3. Balsam LB, Wagers AJ, Christensen JL, et al. Haematopoietic stem cells adopt mature haematopoietic fates in ischemic myocardium. Nature 2004;428:668-673.

4. Follmann D, Wu M. An approximate generalized linear model with random effects for informative missing data. *Biometrics*. 1995 Mar;51(1):151-68.

5. Ford ES, Ajani UA, Croft JB, Critchley JA, Phil D, Sc M, Labarthe DR, Kottke TE, Giles WH, Capewell S: Explaining the decrease in U.S. deaths from coronary disease, 1980-2000. N Engl J Med 2007;356:2388-98.

6. Gehan EA. "A generalized Score test for comparing arbitrarily singly-censored samples". *Biometrika*, 52 203-223 (1965).

7. Jackson KA, Majka SM, Wang H, et al. Regeneration of ischemic cardiac muscle and vascular endothelium by adult stem cells. J Clin Invest 2001;107:1395-1402.

8. Janssens S, Dubois C, Bogaert J, et al. Autologous bone marrow-derived stemcell transfer in patients with ST-segment elevation myocardial infarction: doubleblind, randomised controlled trial. Lancet 2006;367:113-121.

9. Kapoor A, et. al., Ind. J. Radiol.Imag. 2006; 16 (3): 285-292.

10. Kocher AA, Schuster MD, Szabolcs MJ, et al. Neovascularization of ischemic myocardium by human bone-marrow-derived angioblasts prevents cardiomyocyte apoptosis, reduces remodeling and improves cardiac function. Nat Med 2001;7:430-436.

11. Kowalski J, Tu XM. Modern Applied U-statistics. New York. Wiley. 2007.

12. Lachin J. Worst-rank score analysis with informatively missing observations in clinical trials. Controlled Clinical Trials 1999; 20:408– 422.

13. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, Murabito JM, Vasan RS: Long-term trends in the incidence of and survival with heart failure. N Engl J Med 2002;347:1397-402.

14. Lunde K, Solheim S, Aakhus S, et al. Intracoronary injection of mononuclear bone marrow cells in acute wall infarction. NEJM 2006;355:1199-1209.

15. Moyé LA, Davis BR, Hawkins CM. Analysis of a Clinical Trial Involving a Combined Mortality and Adherence Dependent Interval Censored Endpoint. S*tatistics in Medicine*. Vol. 11, 1705-1717 (1992).

16. Moyé LA for the SAVE Cooperative Group. Rationale and Design of a Trial to Assess Patient Survival and Ventricular Enlargement after Myocardial Infarction. *Am J Cardiol* 1991;68:70D-79D.

17. Murry CE, Soopa MH, Reinecke H, et al. Haematopoietic stem cells do not transdifferentiate into cardiac myocytes in myocardial infarcts. Nature 2004;428:664-668.

18. National Institute of Neurological Disorders and Stroke rt- PA Stroke Study Group. (1995). Tissue plasminogen activator for acute ischemic stroke. New England Journal of Medicine 333, 1581-1587.

19. O'Brien PC. Procedures for Comparing Samples with Multiple EndpointsBiometrics,;40; 1079-1087.

20. Orlic D, Kajstura J, Chimenti S, et al. Bone marrow cells regenerate infracted myocardium. Nature 2001;410:701-705.

21. Penicka M, Horak J, Kobylka P, Pytlik R, Kozak T, Belohlavesk O, Lang O,Slakicka H, Simek S, Palacek T, Lanhart A, Aschermann M, Widimsky P (2007). Research Correspondence: Intracornonary injection of autologous bone marrow-derived mononuclear cells in patients with large anterior acute myocardial infarction: A prematurely terminated randomized study. Journal of the American College of Cardiology. 49:2373-2374.

22. Pfeffer MA, Braunwald, E, Moyé LA, Basta L, Brown EJ, Cuddy TE, Davisr BR, Geltman EM, Goldman S, Flaker GC, Klein M, Lamas GA, Packer M, Rouleau J, Rouleau JL, Rutherford J, Wertheimer JH, Hawkins CM on behalf of the SAVE Investigators (1992) Effect of Captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction - results of the Survival and Ventricular Enlargement Trial. *N Eng J Med* 327(10):669-677, Sep 3, 1992.

23. Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, Jacobsen SJ: Trends in heart failure incidence and survival in a community-based population. JAMA 2004;292:344-50.

24. Rosamond W, Fleal K, Friday G, Furie K, Go A, Greenlund K, Haase N; Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell CJ, Roger V, Rumsfeld J, Sorlie P, Steinberger J, Thom T; Wasserthiel-Smoller S, Hong Y,: Heart disease and stroke statistics--2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 2007;115:e69-171.

25. Schächinger V, Erbs S, Elsässer A, et al. Intracoronary bone marrow-derived progenitor cells in acute myocardial infarction. NEJM 2006;355:1210-1221.

26. Strauer BE, Brehm M, Zeus T, et al. Repair of infarcted myocardium by autologous intracoronary mononuclear bone marrow cell transplantation in humans. Circulation 2002;106:1913-1918.

27. Tang, DI, Gnecco C, and Geller NL. (1989a). An approximate likelihood ratio test for a normal mean vector with nonnegative components with application to clinical trials. Biometri'ka 76, 577-583.

28. Tang DI, Gnecco C, and Geller NL. (1989b). Design of group sequential clinical trials with multiple endpoints. Journal of the American Statistical Association 84, 776- 779.

29. Tang DI, Geller NL, and Pocock SJ (1993). On the design and analysis of rando mized clinical trials with multiple endpoints. Biometrics 49, 23-30.

30. Tang DI, Getter NL. (1999)Closed Testing Procedures for Group Sequential Clinical Trials with Multiple Endpoints Biometrics, 55; 1188-

1192.

31. Tilley BC, Marler J, Geller NL, Lu M, Legler J, Brott T, Lyden P, and Grotta J (1996). Use of a global test for multiple outcomes in stroke trials with application to the National Institute of Neurological Disorders and Stroke t-PA stroke trial. Stroke 27, 2136- 2142.

32. Wittkowski KM, Lee E, Nussbaum R, Chamian FN, Krueger JG. Combining several ordinal measures in clinical studies. *Statistics in Medicine*. Vol. 23, 1579-1592 (2004).

33. Wollert KC, Meyer GP, Lotz J, et al. Intracoronary autologous bone-marrow cell transfer after myocardial infarction: the BOOST randomized controlled clinical trial. Lancet 2004;364:141-148.

34. Yoon YS, Wecker A, Heyd L, et al. Clonally expanded novel multipotent stem cells from human bone marrow regenerate myocardium after myocardial infarction. JCI 2005;115:326-338.

Table 1. Values of $\varphi_{ij}\varphi_{ij'}$ used in computing $\boldsymbol{E}\left[\varphi_{ij}\varphi_{ij'}\right]$ for Case 1.

	$C_{x(i)}(+,R) < C_{Y(j)}(+,R)$	$C_{x(i)}(+,R) < C_{Y(j)}(-,R)$	$C_{x(i)}(+,R) > C_{Y(j)}(+,R)$	$C_{x(i)}(-,R) > C_{Y(j)}(+,R)$	$C_{x(i)}(-,R) < C_{Y(j)}(-,R)$	$C_{x(i)}(-,R) > C_{Y(j)}(-,R)$
$C_{x(i)}(+,R) < C_{Y(j)}(+,R)$	1	1	−1			
$C_{x(i)}(+,R) < C_{Y(j)}(-,R)$	1	1	−1			
$C_{x(i)}(+,R) > C_{Y(j)}(+,R)$	−1	−1	1			
$C_{x(i)}(-,R) > C_{Y(j)}(+,R)$				1	$-c$	c
$C_{x(i)}(-,R) < C_{Y(j)}(-,R)$				$-c$	c^2	$-c^2$
$C_{x(i)}(-,R) > C_{Y(j)}(-,R)$				c	$-c^2$	c^2

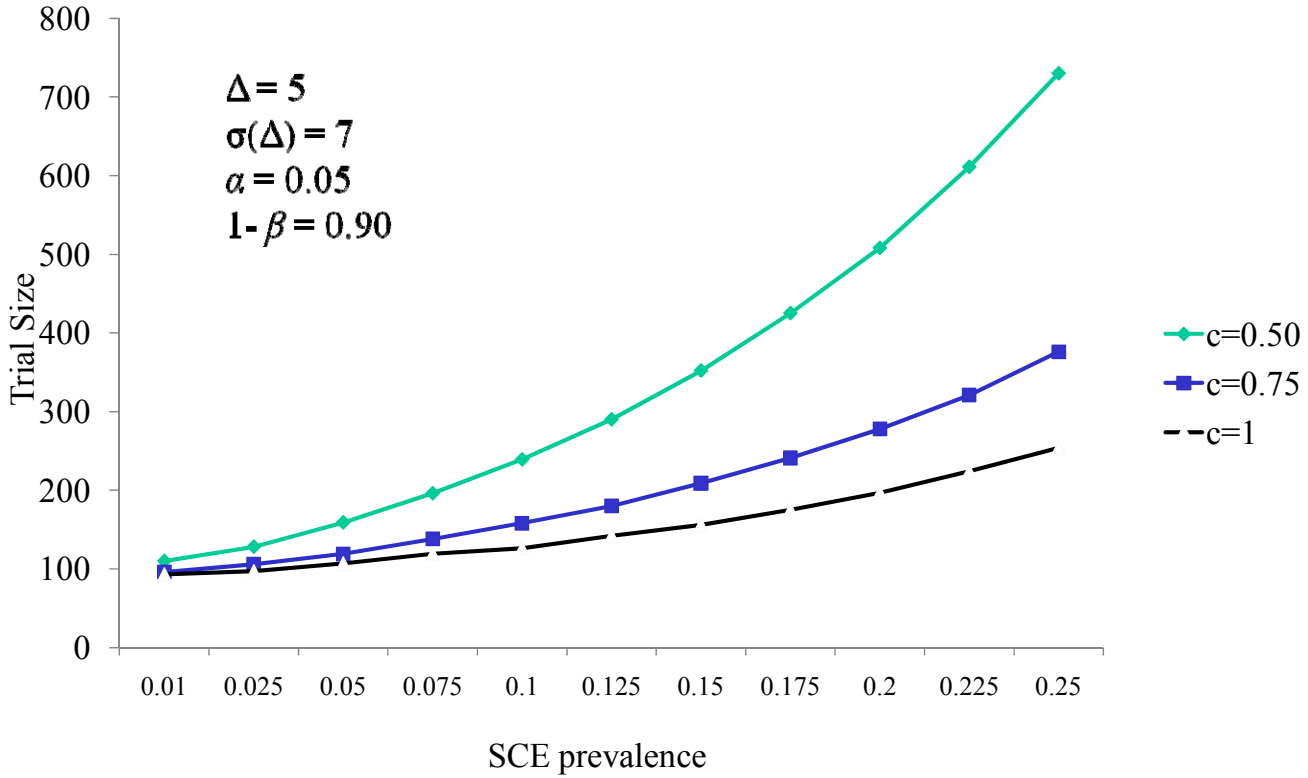

Figure 1. Continuous endpoint weight c influence on the relationship between SCE prevalence and sample size

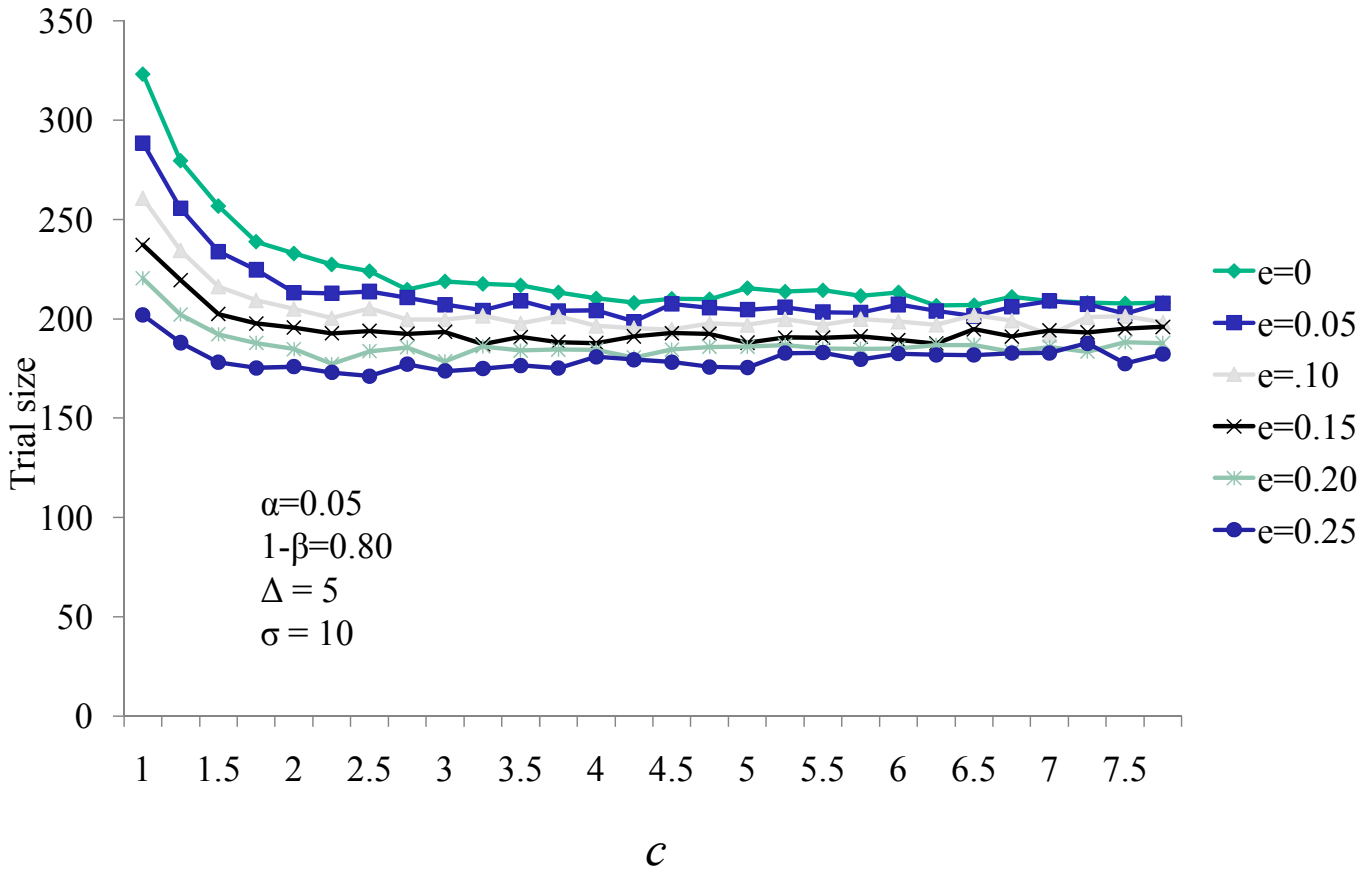

Figure 2. Influence of efficacy on the relationship between the continuous endpoint weight *c* and trial size

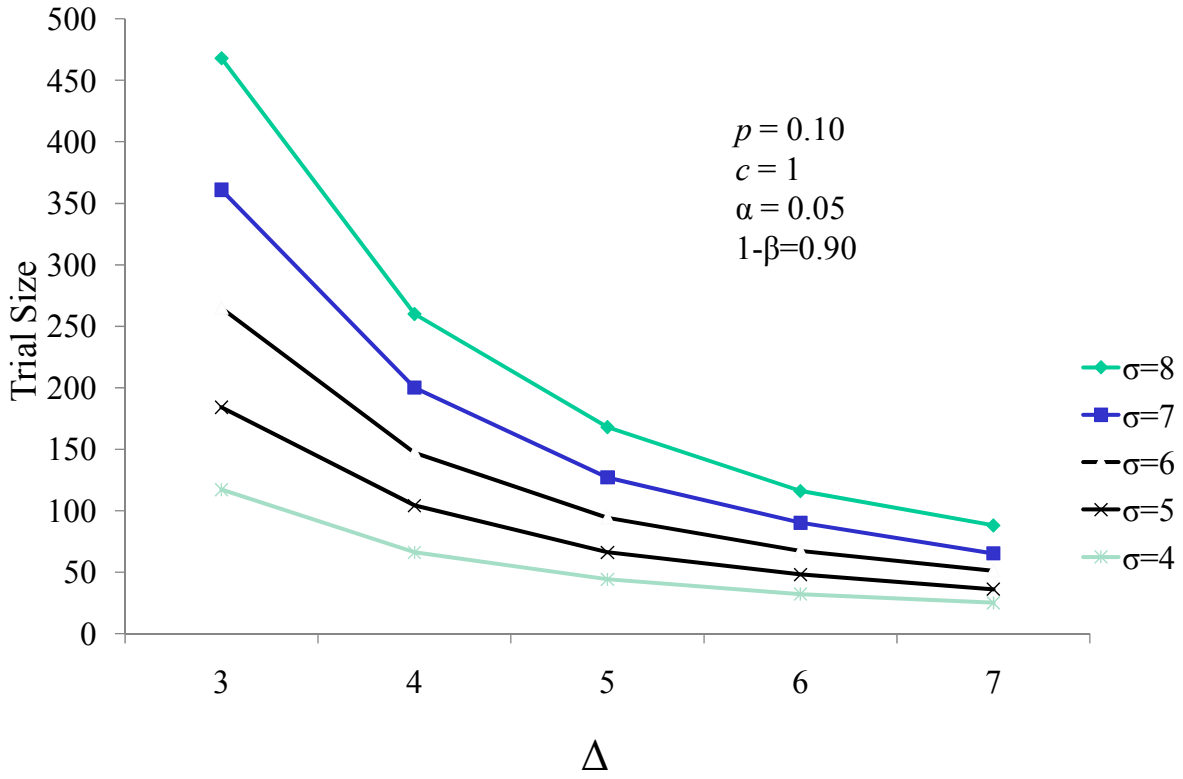

Figure 3. Trial size as a function of standard deviation (σ) and effect size (Δ) of the response variable.

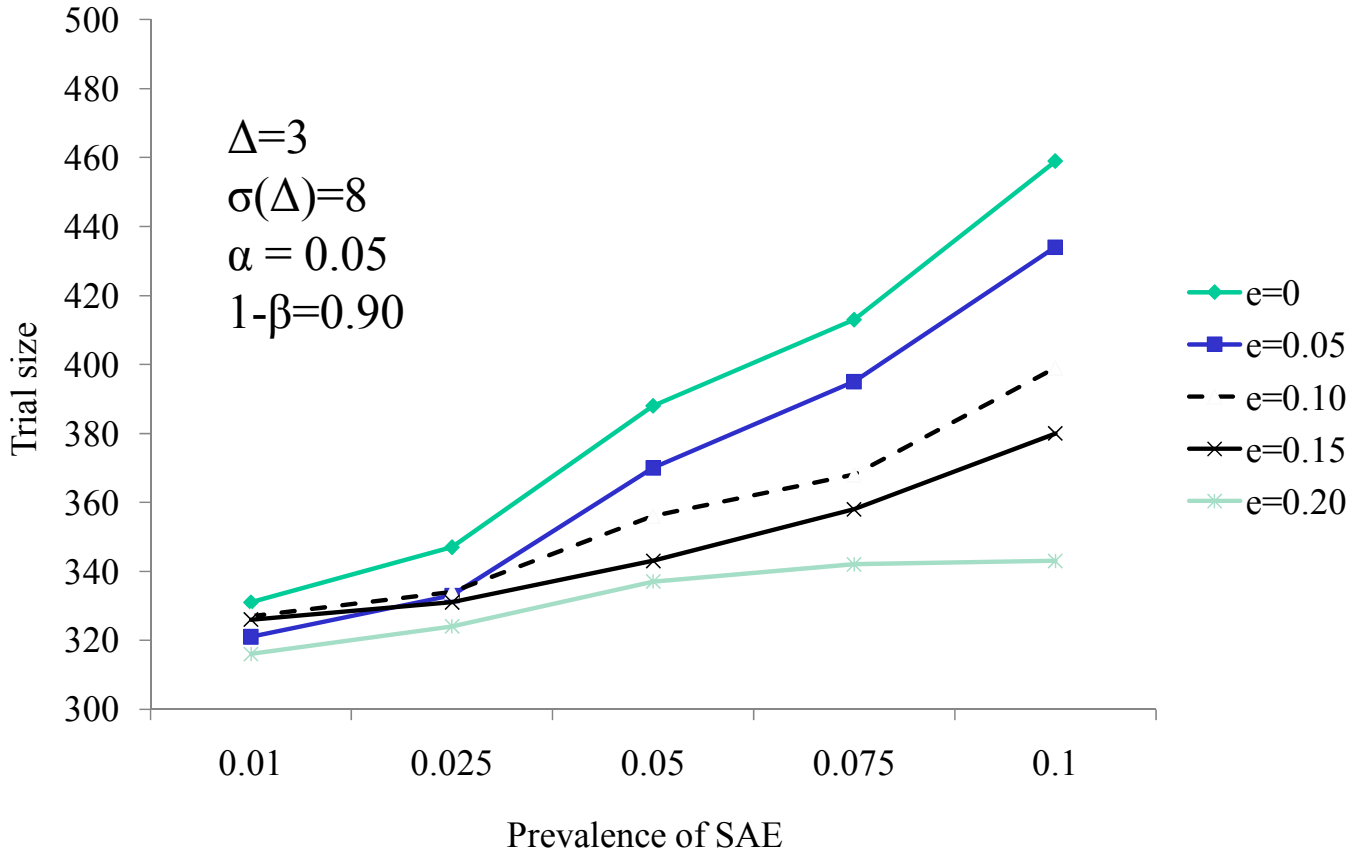

Figure 4. Trial size as a function of the SCE prevalence (*p*) and the intervention's efficacy on prevalence (*e*)

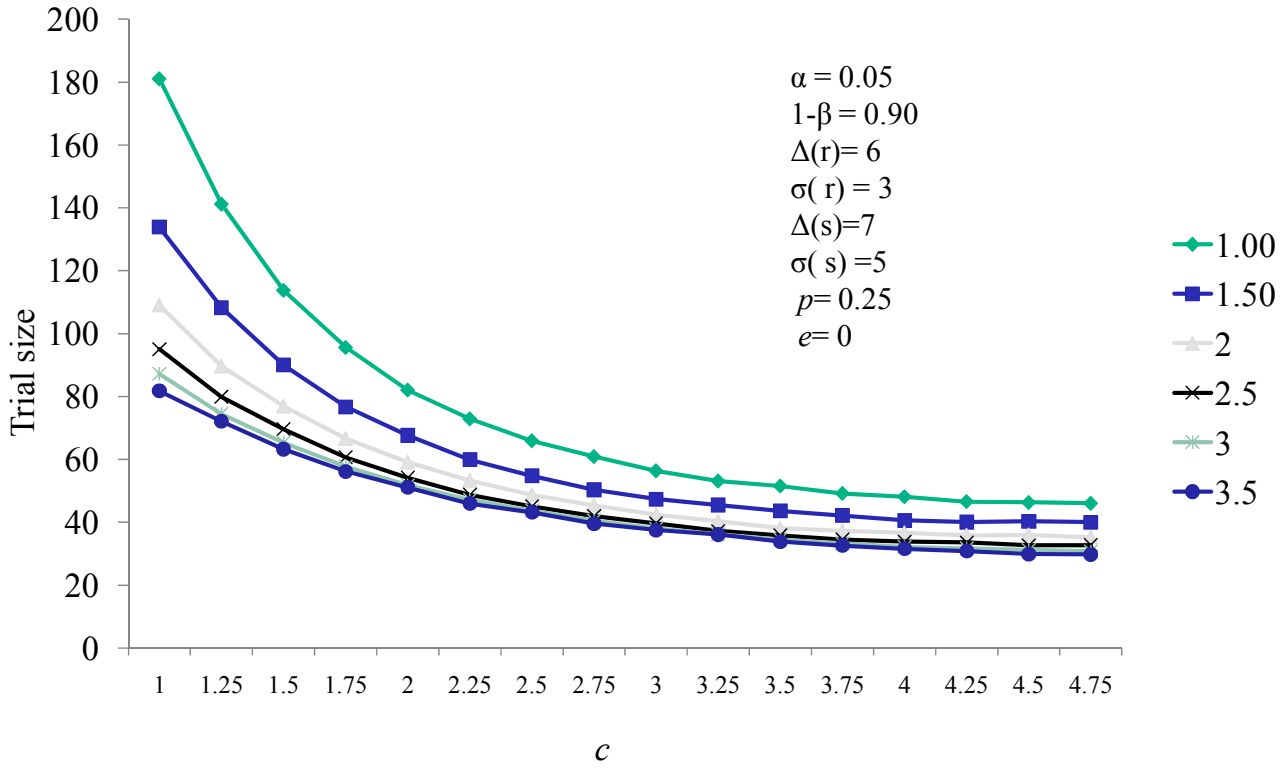

Figure 5. Relationship between sample sizes and weighting factors *c* and *d*.