Analysis of multicentre clinical trials with binary response

By VALERII FEDOROV

Biomedical Data Sciences, GlaxoSmithKline Pharmaceuticals, PO Box 5089, Collegeville PA 19426-0989, USA

Valeri.V.Fedorov@gsk.com

BYRON JONES

Statistical Research and Consulting Centre, Pfizer Global Research and Development, Ramsgate Road, Sandwich, Kent, CT13 9NJ, UK

Byron.Jones@pfizer.com

VIPPAL SAVANI AND ANATOLY ZHIGLJAVSKY[1]

School of Mathematics, Cardiff University, Cardiff, CF24 4AG, UK

SavaniV@cardiff.ac.uk, ZhigljavskyAA@cardiff.ac.uk

SUMMARY

The design and analysis of multicentre trials based on a random effects model is well developed for a continuous response, but is less well developed for a binary response. Here we describe a random effects model for a binary response for two treatments and show how maximum likelihood estimates for the unknown treatment difference can be derived using an approximation to the likelihood. From these results we develop an expression for the Fisher information matrix of the treatment parameters and show how this can be used to determine an optimal design for a multicentre trial for two treatments. The results extend those previously reviewed by Agresti and Hartzel (2000).

Key words: multicentre clinical trials; binary response; maximum likelihood; asymptotic variance; estimating treatment effect.

[1]Corresponding author

1. Introduction

In a multicentre trial patients are enrolled at multiple centres and the patient responses from these centres are combined to give a single estimate of the treatment difference. When the responses are normally distributed, methods for combining the results from multiple centres are well understood and there is a large and growing literature on the topic. For a review of these results see Fedorov and Jones (2005), for example. They describe and compare the well-known fixed effects models and present a random effects model for the case where the treatment effects are a random sample from a population of centres. They also give a brief account of a quasi-linear approximate method for the analysis of binary data using a random effects model. Here we adopt a more conventional maximum likelihood approach, which extends the results reviewed by Agresti and Hartzel (2000). In Section 2 we introduce our notation and the random effects model. In Section 3 we define the likelihood for binary multicentre data and develop an approximation to the required likelihood that avoids the use of numerical integration. This is done using approximating integrals, which is shown to be very accurate. Having developed this approximation we derive expressions for the estimates of the mean response under each treatment. In Section 4 we derive an approximation for the Fisher information matrix of the treatment mean estimators. Particular difficulties that arise when there are no observed successes or failures in a centre are considered in Section 5; solutions to overcome these are derived.

2. Random effects models for multicentre data

Let y_{ijk} denote the binary response observed on patient k, who received treatment j in centre i, where $i = 1, 2, \ldots, N$; $j = 1, 2$ and $k = 1, 2, \ldots, n_{ij}$. If the response is a success, $y_{ijk} = 1$ and if it is a failure $y_{ijk} = 2$. Let r_{ijq} denote the number of patients on treatment j in centre i, that give the response $q = 1$ or 2; we have $n_{ij} = r_{ij1} + r_{ij2}$.

Let π_{ij} denote the probability of a success for treatment j in centre i. We assume that all the responses in centre i are independent and identically distributed (i.i.d.) Bernoulli random variables with success probabilities π_{ij}. This implies that r_{ijq} has a Binomial distribution with parameters (n_{ij}, π_{ij}).

We will model the success probabilities on the logit scale, where $\text{logit}(\pi_{ij}) = \log[\pi_{ij}/(1-\pi_{ij})]$:

$$\begin{aligned} \mu_{i1} = \text{logit}(\pi_{i1}) &= \mu_i + \frac{\delta_i}{2} \\ \mu_{i2} = \text{logit}(\pi_{i2}) &= \mu_i - \frac{\delta_i}{2}\,. \end{aligned}$$

Here μ_i is the effect of centre i and δ_i is the difference between the two treatment effects in centre i, i.e., $\delta_i = \mu_{i1} - \mu_{i2}$. For the success probabilities π_{ij} $(j = 1, 2)$ we have the following expressions:

$$\begin{aligned} \pi_{i1} &= \frac{\exp(\mu_{i1})}{1+\exp(\mu_{i1})} = \frac{\exp\left(\mu_i + \frac{\delta_i}{2}\right)}{1+\exp\left(\mu_i + \frac{\delta_i}{2}\right)}, \\ \pi_{i2} &= \frac{\exp(\mu_{i2})}{1+\exp(\mu_{i2})} = \frac{\exp\left(\mu_i - \frac{\delta_i}{2}\right)}{1+\exp\left(\mu_i - \frac{\delta_i}{2}\right)}. \end{aligned}$$

This model allows a unique treatment difference to occur in each centre, i.e., it allows for treatment-by-centre interaction. In the random effects setting, μ_{i1} and μ_{i2} are considered as random variables having some bivariate distribution. We shall assume that the bivariate distribution is normal $\mathbb{N}(\boldsymbol{\nu}, \mathbb{V})$ with unknown mean $\boldsymbol{\nu} = (\nu_1, \nu_2)^T$ and known variance-covariance matrix

$$\mathbb{V} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

That is, we assume that the vectors $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2})^T$, $(i = 1, 2, \ldots, N)$ are random, independent and follow the bivariate normal distribution with density

$$f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(\mu_{i1}-\nu_1)^2}{\sigma_1^2} - 2\rho\frac{(\mu_{i1}-\nu_1)(\mu_{i2}-\nu_2)}{\sigma_1\sigma_2} + \frac{(\mu_{i2}-\nu_2)^2}{\sigma_2^2}\right]\right\}.$$

As the estimation of the treatment difference is the main aim of the trial, primary interest focuses on the expectation and variance of δ_i. To obtain these we note that

$$\mu_i = \frac{\mu_{i1} + \mu_{i2}}{2} \quad \text{and} \quad \delta_i = \mu_{i1} - \mu_{i2}.$$

The joint distribution of μ_i and δ_i is given in the following lemma (the proof of this lemma is straightforward and therefore omitted).

Lemma 1. If

$$\begin{pmatrix} \mu_{i1} \\ \mu_{i2} \end{pmatrix} \sim \mathbb{N}\left[\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right],$$

then

$$\begin{pmatrix} \mu_i \\ \delta_i \end{pmatrix} \sim \mathbb{N}\left[\begin{pmatrix} \nu_0 \\ \delta \end{pmatrix}, \begin{pmatrix} \frac{1}{4}(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2) & \frac{1}{2}(\sigma_1^2 - \sigma_2^2) \\ \frac{1}{2}(\sigma_1^2 - \sigma_2^2) & \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \end{pmatrix}\right],$$

where $\nu_0 = (\nu_1 + \nu_2)/2$ and $\delta = \nu_1 - \nu_2$.

3. Estimating the centre and treatment effects

3.1. *Approximating the likelihood function.*

To estimate the unknown parameters $\boldsymbol{\nu} = (\nu_1, \nu_2)^T$ we use the following likelihood function, where integration is over the joint distribution of $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2})^T$:

$$l(\boldsymbol{\nu}, \mathbb{V}) = \prod_{i=1}^{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \prod_{j=1}^{2} (\pi_{ij})^{r_{ij1}} (1-\pi_{ij})^{r_{ij2}} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \,, \tag{1}$$

where $d\boldsymbol{\mu}_i = d\mu_{i1} d\mu_{i2}$.

The expression (1) for the likelihood function generalizes (to the case of general covariance matrix $\mathbb{V}$) the formula given on page 1120 of Agresti and Hartzel (2000), but note that the product over j (which is k in their expression) should have been moved inside the integrals.

In order to evaluate the likelihood we must integrate the joint mass function of the responses with respect to the random effects distribution. A common approach to this is to approximate the likelihood function using numerical integration methods, such as Gauss-Hermit quadrature, as discussed by Agresti and Hartzel (2000) and used in SAS® (Statistical Analysis System, SAS Institute Inc., Cary, NC, USA) as a standard. We also approximate the integrals but use a different approach based on approximating the intergrands in such a way that the integrals can be easily evaluated. This allows us to obtain good approximations and explicit formulas for both estimators and their asymptotic variances.

Let us rewrite the likelihood function as $l(\boldsymbol{\nu}, \mathbb{V}) = \prod_{i=1}^{N} I_i$, where

$$\begin{aligned} I_i &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{e^{\mu_{i1}}}{1+e^{\mu_{i1}}}\right]^{r_{i11}} \left[1-\frac{e^{\mu_{i1}}}{1+e^{\mu_{i1}}}\right]^{r_{i12}} \left[\frac{e^{\mu_{i2}}}{1+e^{\mu_{i2}}}\right]^{r_{i21}} \left[1-\frac{e^{\mu_{i2}}}{1+e^{\mu_{i2}}}\right]^{r_{i22}} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp\left(\mu_{i1} r_{i11}\right)}{\left[1+\exp(\mu_{i1})\right]^{(r_{i11}+r_{i12})}} \frac{\exp\left(\mu_{i2} r_{i21}\right)}{\left[1+\exp(\mu_{i2})\right]^{(r_{i21}+r_{i22})}} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \,. \end{aligned} \tag{2}$$

Let us introduce the function

$$\phi(t; x, y) = \frac{\exp(xt)}{[1+\exp(t)]^{x+y}} \,, \tag{3}$$

where $x \geq 0$ and $y \geq 0$ are parameters and t is the variable. Then we can write

$$I_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\mu_{i1}; r_{i11}, r_{i12})\, \phi(\mu_{i2}; r_{i21}, r_{i22}) f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \,. \tag{4}$$

Assume that $x > 0$ and $y > 0$. The function (3) achieves its maximum value at the point $t_* = \log(x/y)$. Expanding the logarithm of the function (3) around the point t_* we obtain

after some simplifications:

$$\log \phi(t;x,y) = \log\left[\frac{x^x y^y}{(x+y)^{x+y}}\right] - \frac{xy(t-\log(x/y))^2}{2(x+y)} + O\left(|\, t-\log(x/y)\,|^3\right) .$$

This gives us the following approximation for the function (3):

$$\phi(t;x,y) \simeq \psi(t;x,y) = \left[\frac{x^x y^y}{(x+y)^{x+y}}\right] \exp\left\{-\frac{xy(t-\log(x/y))^2}{2(x+y)}\right\} . \tag{5}$$

If x and y are not too small, the quality of approximation (5) is very good; Figure 1 illustrates this in the case $x = 3, y = 4$. The fact that the tails of the functions $\phi(t;x,y)$ and $\psi(t;x,y)$ behave differently is of no importance in the present circumstances as the functions are to be multiplied by a normal density and then integrated.

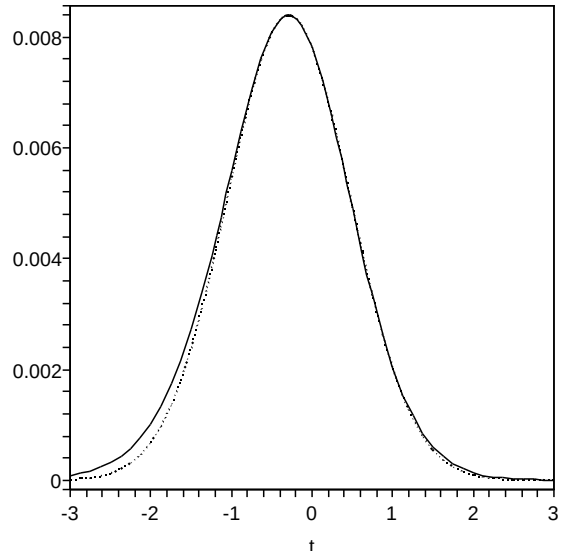

FIGURE 1. Graphs of the functions $\phi(t;x,y)$ (solid line) and $\psi(t;x,y)$ (dash-dot) for $x = 3$ and $y = 4$

Thus, we can approximate the components of the likelihood function as follows:

$$I_i \simeq J_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(r_{i11})^{r_{i11}}(r_{i12})^{r_{i12}}}{(r_{i11}+r_{i12})^{r_{i11}+r_{i12}}} \exp\left\{-\frac{1}{2}\frac{r_{i11}r_{i12}\left(\mu_{i1}-\log\left(\frac{r_{i11}}{r_{i12}}\right)\right)^2}{r_{i11}+r_{i12}}\right\}$$

$$\times \frac{(r_{i21})^{r_{i21}}(r_{i22})^{r_{i22}}}{(r_{i21}+r_{i22})^{r_{i21}+r_{i22}}} \exp\left\{-\frac{1}{2}\frac{r_{i21}r_{i22}\left(\mu_{i2}-\log\left(\frac{r_{i21}}{r_{i22}}\right)\right)^2}{r_{i21}+r_{i22}}\right\} f(\boldsymbol{\mu}_i;\boldsymbol{\nu},\mathbb{V})d\boldsymbol{\mu}_i$$

$$= A_i \frac{1}{2\pi s_{i1}s_{i2}} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-\frac{(\mu_{i1}-\gamma_{i1})^2}{2s_{i1}^2} - \frac{(\mu_{i2}-\gamma_{i2})^2}{2s_{i2}^2}\right\} \times$$

$$\exp\left\{-\frac{(\mu_{i1}-\nu_1)^2}{2\sigma_1^2(1-\rho^2)} - \frac{(\mu_{i2}-\nu_2)^2}{2\sigma_2^2(1-\rho^2)} + \frac{\rho(\mu_{i1}-\nu_1)(\mu_{i2}-\nu_2)}{\sigma_1\sigma_2(1-\rho^2)}\right\} d\boldsymbol{\mu}_i\,,$$

where

$$s_{i1} = \sqrt{\frac{r_{i11}+r_{i12}}{r_{i11}r_{i12}}}\,, \quad s_{i2} = \sqrt{\frac{r_{i21}+r_{i22}}{r_{i21}r_{i22}}}\,, \quad \gamma_{i1} = \log\left(\frac{r_{i11}}{r_{i12}}\right), \quad \gamma_{i2} = \log\left(\frac{r_{i21}}{r_{i22}}\right)$$

and $A_i = C_i(r_{i11}, r_{i21})$; for $0 < x < n_{i1}$ and $0 < y < n_{i2}$ the function $C_i(x,y)$ is defined as

$$C_i(x,y) = 2\pi \frac{x^{x-1/2}y^{y-1/2}(n_{i1}-x)^{n_{i1}-x-1/2}(n_{i2}-y)^{n_{i2}-y-1/2}}{(n_{i1})^{n_{i1}-1/2}(n_{i2})^{n_{i2}-1/2}}\,. \tag{6}$$

After integration we obtain the following expression for J_i:

$$J_i = A_i \frac{1}{2\pi\sqrt{(s_{i1}^2+\sigma_1^2)(s_{i2}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2}}$$

$$\times \exp\left\{-\frac{(s_{i2}^2+\sigma_2^2)(\nu_1-\gamma_{i1})^2 - 2\rho\sigma_1\sigma_2(\nu_1-\gamma_{i1})(\nu_2-\gamma_{i2}) + (s_{i1}^2+\sigma_1^2)(\nu_2-\gamma_{i2})^2}{2\left[(s_{i1}^2+\sigma_1^2)(s_{i2}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2\right]}\right\}.$$

The expression for J_i/A_i is in fact a value of a density function of a suitable bivariate normal distribution taken at $\boldsymbol{\nu} = (\nu_1,\nu_2)^T$; that is, $J_i = A_i f(\boldsymbol{\nu};\boldsymbol{\gamma}_i,\tilde{\mathbb{V}}_i)$ with

$$\boldsymbol{\gamma}_i = \begin{pmatrix}\gamma_{i1}\\ \gamma_{i2}\end{pmatrix} \quad \text{and} \quad \tilde{\mathbb{V}}_i = \begin{pmatrix}\tilde{\sigma}_{i1}^2 & \tilde{\rho}\tilde{\sigma}_{i1}\tilde{\sigma}_{i2}\\ \tilde{\rho}_i\tilde{\sigma}_{i1}\tilde{\sigma}_{i2} & \tilde{\sigma}_{i2}^2\end{pmatrix},$$

where

$$\tilde{\sigma}_{i1} = \sqrt{s_{i1}^2+\sigma_1^2}, \;\; \tilde{\sigma}_{i2} = \sqrt{s_{i2}^2+\sigma_2^2}, \;\; \text{and} \;\; \tilde{\rho}_i = \rho\frac{\sigma_1\,\sigma_2}{\tilde{\sigma}_{i1}\,\tilde{\sigma}_{i2}}\,.$$

3.2. *Computing the maximum likelihood estimates.*

Taking the logarithm of our likelihood function (1) and using the properties of the log function we have

$$\log l \simeq \sum_{i=1}^{N} \log(J_i) = Q(\nu_1, \nu_2)\,. \tag{7}$$

Here Q is a quadratic function in ν_1 and ν_2:

$$Q(\nu_1, \nu_2) = -\frac{1}{2}K_1\nu_1^2 - \frac{1}{2}K_2\nu_2^2 + K_3\nu_1\nu_2 + K_4\nu_1 + K_5\nu_2 + K_6\,,$$

where $K_j = \sum_{i=1}^{N} K_{ji}$ with

$$\begin{aligned}
K_{1i} &= \frac{1}{(1-\tilde{\rho_i}^2)\tilde{\sigma}_{i1}^2} = \frac{s_{i2}^2+\sigma_2^2}{(s_{i1}^2+\sigma_1^2)(s_{i2}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2},\\
K_{2i} &= \frac{1}{(1-\tilde{\rho_i}^2)\tilde{\sigma}_{i2}^2} = \frac{s_{i1}^2+\sigma_1^2}{(s_{i1}^2+\sigma_1^2)(s_{i2}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2},\\
K_{3i} &= (\rho\sigma_1\sigma_2)^{-1}\frac{\tilde{\rho_i}^2}{(1-\tilde{\rho_i}^2)} = \frac{\rho\sigma_1\sigma_2}{(s_{i1}^2+\sigma_1^2)(s_{i2}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2},\\
K_{4i} &= \gamma_{i1}K_{1i} - \gamma_{i2}K_{3i}, \quad K_{5i} = \gamma_{i2}K_{2i} - \gamma_{i1}K_{3i}, \quad \text{and}\\
K_{6i} &= -\frac{1}{2}\gamma_{i1}^2K_{1i} + \gamma_{i1}\gamma_{i2}K_{3i} - \frac{1}{2}\gamma_{i2}^2K_{2i} + \log(A_i) - \log\left(2\pi\tilde{\sigma}_{i1}\tilde{\sigma}_{i2}\sqrt{1-\tilde{\rho_i}^2}\right).
\end{aligned}$$

Differentiating Q with respect to ν_1 and ν_2 and equating the derivatives to zero we get

$$\frac{\partial Q}{\partial \nu_1} = -K_1\nu_1 + K_3\nu_2 + K_4 = 0$$

and

$$\frac{\partial Q}{\partial \nu_2} = -K_2\nu_2 + K_3\nu_1 + K_5 = 0\,.$$

Solving these equations we obtain

$$\hat{\nu}_1 = \frac{K_3K_5 + K_2K_4}{K_1K_2 - K_3^2} \;\text{ and }\; \hat{\nu}_2 = \frac{K_1K_5 + K_4K_3}{K_1K_2 - K_3^2}. \tag{8}$$

We note that these estimators can be evaluated directly without the need to apply iterative procedures. This is an important advantage because in the traditional approach iterative procedures are needed to compute the maximum likelihood estimators (e.g., see Agresti and Hartzel, 2000). Some results of comparison of the estimators (8) with SAS® estimators are provided in Table 1 and Figure 2 (To compute the SAS® estimators, we have used exactly the same code as given in Agresti and Hartzel, 2000). The general conclusion is that the estimators (8) and the estimators produced by SAS® are very similar, with the estimators (8) often having slightly smaller MSE's.

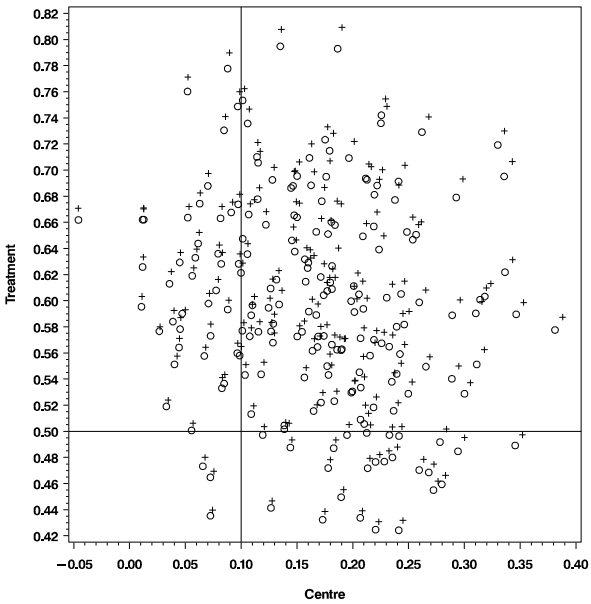

FIGURE 2. A sample of 200 estimators computed using (8) (depicted as circles) and adaptive Gaussian quadratures in SAS® (crosses) in the case of 50 centres, 100 patients on each arm, centre effect $\nu_0 = 0.1$, treatment effect $\delta = 0.5$, $\sigma_1^2 = \sigma_2^2 = \frac{3}{4}$ and $\rho = \frac{1}{3}$.

In multicentre trials it is not unusual in one or more centres, for all the patients to respond with a success or all patients to respond with a failure. As the above estimators cannot be computed if this happens, some adjustments are needed. These are described in Section 5.

4. Estimating the variance of the treatment parameters

To estimate the variance-covariance matrix of the estimated treatment parameters we may use the Cramer-Rao theorem, which states that for a random sample of n observations the maximum likelihood estimators are asymptotically normal as $n \to \infty$ with the normalised variance-covariance matrix equal to the inverse of Fisher's information matrix given by

$$I(\boldsymbol{\nu}) = \left[-E\left(\frac{\partial^2 \log l}{\partial \nu_i \partial \nu_j} \right) \right]_{i,j=1,2}.$$

Using the approximation (7) for the loglikelihood function, we obtain the following approximation for the Fisher information matrix:

$$I(\boldsymbol{\nu}) \simeq \left[-E\left(\frac{\partial^2 \log Q(\nu_1, \nu_2)}{\partial \nu_i \partial \nu_j} \right) \right]_{i,j=1,2} = E\left[\begin{pmatrix} K_1 & -K_3 \\ -K_3 & K_2 \end{pmatrix} \right] = \sum_{i=1}^{N} E\left[\begin{pmatrix} K_{1i} & -K_{3i} \\ -K_{3i} & K_{2i} \end{pmatrix} \right].$$

In order to derive expressions for the expected values given in the above formula, we first consider the special case of a trial with a single centre. The results obtained will then be generalized to the multicentre setting.

4.1. *Single Centre Problem.*

Let us consider the problem of estimating the variance-covariance matrix of the estimators (8) for a single centre, and thus we assume that $N = 1$ and $i = 1$. We shall retain the standard notation to avoid confusion later on when considering multiple centres.

Consider the problem of approximating the expectation of K_{1i} which requires taking expectations with respect to r_{i11} and r_{i12}. Since r_{i11} and r_{i12} have a distribution with random parameters we need to take a double expectation, first with respect to the distribution of r_{i11} and r_{i12} and then with respect to the distribution of (μ_{i1}, μ_{i2}).

The expectation is then given by

$$E\left[K_{1i}\right] \;=\; E\left[\sum_{x=0}^{n_{i1}}\sum_{y=0}^{n_{i2}} K_{1i}(x,y)\binom{n_{i1}}{x}\pi_{i1}^{x}(1-\pi_{i1})^{n_{i1}-x}\binom{n_{i2}}{y}\pi_{i2}^{y}(1-\pi_{i2})^{n_{i2}-y}\right],$$

where

$$K_{1i}(x,y) = \frac{s_{iy}^2+\sigma_2^2}{(s_{ix}^2+\sigma_1^2)(s_{iy}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2}$$

with

$$s_{ix}^2 = \frac{n_{i1}}{x(n_{i1}-x)} \quad \text{and} \quad s_{iy}^2 = \frac{n_{i2}}{y(n_{i2}-y)}\,.$$

Replacing the order of summation and integration, we obtain

$$K_{1i}(x,y) = \sum_{x=0}^{n_{i1}}\sum_{y=0}^{n_{i2}} K_{1i}(x,y)\binom{n_{i1}}{x}\binom{n_{i2}}{y} I_i(x,y)\,,$$

where

$$\begin{aligned} I_i(x,y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp\left(\mu_{i1}x\right)}{\left[1+\exp(\mu_{i1})\right]^{(n_{i1}-x)}} \frac{\exp\left(\mu_{i2}y\right)}{\left[1+\exp(\mu_{i2})\right]^{(n_{i2}-y)}} f(\boldsymbol{\mu}_i;\boldsymbol{\nu},\mathbb{V})d\boldsymbol{\mu}_i \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\mu_{i1};x,n_{i1}-x)\,\phi(\mu_{i2};y,n_{i2}-y) f(\boldsymbol{\mu}_i;\boldsymbol{\nu},\mathbb{V})d\boldsymbol{\mu}_i\,. \end{aligned}$$

The integral $I_i(x,y)$ is the integral of the same type as (4) and can be approximated in the same way. Assuming that $0 < x < n_{i1}$ and $0 < y < n_{i2}$ we obtain the following approximation for $I_i(x,y)$:

$$I_i(x,y) \simeq J_i(x,y) \;=\; C_i(x,y)\frac{1}{2\pi\sqrt{(s_{ix}^2+\sigma_1^2)}\sqrt{(s_{iy}^2+\sigma_2^2)}\sqrt{1-\tilde{\rho}_{i,x,y}^2}}\times$$

$$\exp\left\{-\frac{1}{2\left(1-\tilde{\rho}_{i,x,y}^2\right)}\left[\left(\frac{\nu_1-\gamma_{ix}}{\sqrt{(s_{ix}^2+\sigma_1^2)}}\right)^2-2\tilde{\rho}_{i,x,y}\left(\frac{\nu_1-\gamma_{ix}}{\sqrt{s_{ix}^2+\sigma_1^2}}\right)\left(\frac{\nu_2-\gamma_{iy}}{\sqrt{s_{iy}^2+\sigma_2^2}}\right)+\left(\frac{\nu_2-\gamma_{iy}}{\sqrt{(s_{iy}^2+\sigma_2^2)}}\right)^2\right]\right\}$$

where

$$\gamma_{ix} = \log\left(\frac{x}{n_{i1}-x}\right), \quad \gamma_{iy}=\log\left(\frac{y}{n_{i2}-y}\right), \quad \tilde{\rho}_{i,x,y}=\rho\frac{\sigma_1}{\sqrt{(s_{ix}^2+\sigma_1^2)}}\frac{\sigma_2}{\sqrt{(s_{iy}^2+\sigma_2^2)}}$$

and $C_i(x,y)$ is as defined in (6).

Therefore, if we ignore the terms corresponding to $x=0,\ x=n_{i1},\ y=0$ and $y=n_{i2}$ then we obtain the following approximation for $E\left[K_{1i}\right]$:

$$E\left[K_{1i}\right] \simeq \sum_{x=1}^{n_{i1}-1}\sum_{y=1}^{n_{i2}-1} K_{1i}(x,y)\binom{n_{i1}}{x}\binom{n_{i2}}{y}J_i(x,y)\,.$$

The correction related to the terms $x=0,\ x=n_{i1},\ y=0$ and $y=n_{i2}$ is considered in Section 5.

The variance-covariance matrix of the estimators can now be obtained by deriving the inverse of the Fisher information matrix given by

$$\begin{bmatrix} Var[\hat{\nu_1}] & Cov[\hat{\nu_1},\hat{\nu_2}] \\ Cov[\hat{\nu_1},\hat{\nu_2}] & Var[\hat{\nu_2}] \end{bmatrix} \simeq \frac{1}{E[K_{1i}]E[K_{2i}]-E[K_{3i}]^2}\begin{bmatrix} E[K_{2i}] & E[K_{3i}] \\ E[K_{3i}] & E[K_{1i}] \end{bmatrix}$$

with entries

$$\begin{aligned} E\left[K_{1i}\right] &= \sum_{x=1}^{n_{i1}-1}\sum_{y=1}^{n_{i2}-1} K_{1i}(x,y)\binom{n_{i1}}{x}\binom{n_{i2}}{y}J_i(x,y), \\ E\left[K_{2i}\right] &= \sum_{x=1}^{n_{i1}-1}\sum_{y=1}^{n_{i2}-1} K_{2i}(x,y)\binom{n_{i1}}{x}\binom{n_{i2}}{y}J_i(x,y) \\ \text{and} \quad E\left[K_{3i}\right] &= \sum_{x=1}^{n_{i1}-1}\sum_{y=1}^{n_{i2}-1} K_{3i}(x,y)\binom{n_{i1}}{x}\binom{n_{i2}}{y}J_i(x,y) \end{aligned}$$

where

$$K_{2i}(x,y)=\frac{s_{ix}^2+\sigma_1^2}{(s_{ix}^2+\sigma_1^2)(s_{iy}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2} \ \text{ and } \ K_{3i}(x,y)=\frac{\rho\sigma_1\sigma_2}{(s_{ix}^2+\sigma_1^2)(s_{iy}^2+\sigma_2^2)-\sigma_1^2\sigma_2^2\rho^2}\,.$$

4.2. *Multiple Centre Problem.*

The variance-covariance matrix of the estimators $(\hat{\nu_1}, \hat{\nu_2})$ for the multiple centre problem can be obtained directly from the results of the single centre problem. Taking the inverse of the Fisher information matrix we get

$$\begin{bmatrix} Var[\hat{\nu_1}] & Cov[\hat{\nu_1}, \hat{\nu_2}] \\ Cov[\hat{\nu_1}, \hat{\nu_2}] & Var[\hat{\nu_2}] \end{bmatrix} \approx \frac{1}{\sum_{i=1}^{N} EK_{1i} \sum_{i=1}^{N} EK_{2i} - \left(\sum_{i=1}^{N} EK_{3i}\right)^2} \sum_{i=1}^{N} \begin{bmatrix} EK_{2i} & EK_{3i} \\ EK_{3i} & EK_{1i} \end{bmatrix}.$$

Note that if the number of subjects on treatment 1, (n_{i1}), and treatment 2, (n_{i2}), are the same for each centre $(i = 1, 2, \ldots, N)$ then the variance-covariance matrix is equivalent to the variance-covariance matrix for the single centre estimates reduced by a factor of N, i.e.,

$$\begin{bmatrix} Var[\hat{\nu_1}] & Cov[\hat{\nu_1}, \hat{\nu_2}] \\ Cov[\hat{\nu_1}, \hat{\nu_2}] & Var[\hat{\nu_2}] \end{bmatrix} \approx \frac{1}{N(EK_{1i}\, EK_{2i} - [EK_{3i}]^2)} \begin{bmatrix} EK_{2i} & EK_{3i} \\ EK_{3i} & EK_{1i} \end{bmatrix} \quad \text{with } i = 1.$$

In most cases that we have studied numerically (an example is given in Table 2), the approximations derived above give very similar results to the results that one gets from SAS® using the adaptive Gaussian quadratures (which is the recommended way of estimating integrals of this type). The advantage of our approximation is therefore the fact that we have it in explicit form.

5. Zero-Term Corrections

The estimators $\hat{\nu}_1$ and $\hat{\nu}_2$ given in equations (8) have the disadvantage that they are not defined when there are no observed successes or no observed failures (i.e., $r_{ijq} = 0$ $j, q \in \{1, 2\}$) for any given centre i. The chance of observing such extremes depends on the number of patients in the centre, n_{ij}, and the success probabilities π_{i1} and π_{i2}. The lower the value of n_{ij} and the closer the success probabilities are to zero or one, the more likely we are to observe $r_{ijq} = 0,\ j, q \in \{1, 2\}$. In Sections 5 and 5 we present two methods of adjusting the approximation when we have centres with such extreme observations. First, in Section 5 we consider a particular case $\rho = 0$.

5.1. *The case* $\rho = 0$.

If the correlation ρ between random effects μ_{i1} and μ_{i2} is zero then the bivariate normal density $f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V})$ is a product of two univariate normal densities:

$$f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) = f(\mu_{i1}; \nu_1, \sigma_1^2)\, f(\mu_{i1}; \nu_2, \sigma_2^2)\,,$$

where $f(t;\nu,\sigma^2) = 1/(\sqrt{2\pi}\sigma)\exp\{-(t-\nu)^2/(2\sigma^2)\}$. Hence the double integrals I_i defined in (2) become the products of two single integrals; that is, I_i may by written as

$$I_i = \int_{-\infty}^{\infty} \frac{e^{\mu_{i1} r_{i11}}}{\left[1+e^{\mu_{i1}}\right]^{r_{i11}+r_{i12}}} f(\mu_{i1};\nu_1,\sigma_1^2)d\mu_{i1} \int_{-\infty}^{\infty} \frac{e^{\mu_{i2} r_{i21}}}{\left[1+e^{\mu_{i2}}\right]^{r_{i21}+r_{i22}}} f(\mu_{i2};\nu_2,\sigma_2^2)d\mu_{i2}\,. \qquad (9)$$

Approximating single integrals in (9) in the case when either one or two of $r_{ijq},\ j,q \in \{1,2\}$) is equal to zero is a relatively easy problem. In the following sections we assume $\rho \neq 0$.

5.2. *Zero-Term Regularization.*

The simplest adjustment that can be made is to replace the zero term in the likelihood with a regularization constant. Here we replace the case where there are no observed successes or failures ($r_{ijq} = 0,\ j,q \in \{1,2\}$) with a small positive constant δ, say $\delta = 0.1$; that is, we set $r_{ijq} = \delta,\ j,q \in \{1,2\}$. Then we can use the approximation (7) to the log-likelihood function which remain quadratic with respect to (ν_1,ν_2) and therefore the estimators $\hat{\nu}_1$ and $\hat{\nu}_2$ have exactly the same form. Numerical study shows that if the values of n_{ij} are not too small, the value chosen for δ is not important, as long as it is small.

5.3. *Replacing mean of a function of a random variable with the function of the mean.*

Another natural approximation is obtained by replacing the mean of a function (3) of a random parameter t (which is one of μ_{ij}) by the function of the mean of t. That is, we are going to use the approximation

$$E\phi(\mu_{ij};r_{ij1},r_{ij2}) \simeq \phi(E\mu_{ij};r_{ij1},r_{ij2}), \quad j \in \{1,2\}\,. \qquad (10)$$

Note, however, that the resulting approximation to the log-likelihood will no longer be quadratic in ν_1 and ν_2 and in Section 5 we shall consider further simplifications leading to retaining the quadratic form of the log-likelihood approximations.

The log-likelihood function may be written as

$$\log l \;=\; \sum_{\substack{i=1:\\ r_{ijq}\notin\{0,n_{ij}\}\\ \forall j,q\in\{1,2\}}}^{N} \log I_i + \sum_{\substack{i=1:\\ r_{ijq}\in\{0,n_{ij}\}\\ \text{any}\, j,q\in\{1,2\}}}^{N} \log I_i\,.$$

Let us look at each possible case separately.

Case 1: Observed zero for one of the treatments.

Consider the case when there are no observed successes for subjects taking treatment 1

($r_{i11} = 0$) but no such extremes for subjects taking treatment 2 ($r_{i21} \notin \{0, n_{i2}\}$). In this case the log-likelihood function is given by

$$I_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{[1+\exp(\mu_{i1})]^{n_{i1}}} \frac{\exp(\mu_{i2} r_{i21})}{[1+\exp(\mu_{i2})]^{(r_{i21}+r_{i22})}} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \,.$$

Using the approximation (10) for the function $\phi(\mu_{i1}; 0, n_{i1})$ we obtain

$$\phi(\mu_{i1}; 0, n_{i1}) = \frac{1}{[1+\exp(\mu_{i1})]^{n_{i1}}} \approx \frac{1}{[1+\exp(\nu_1)]^{n_{i1}}}.$$

The integral I_i may then be approximated by

$$\begin{aligned} I_i &\approx \frac{1}{[1+\exp(\nu_1)]^{n_{i1}}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(\mu_{i2} r_{i21})}{[1+\exp(\mu_{i2})]^{(r_{i21}+r_{i22})}} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i \\ &\approx b_i \frac{1}{(1+\exp(\nu_1))^{n_{i1}}} \frac{1}{\sqrt{2\pi(s_{i2}^2 + \sigma_2^2)}} \exp\left\{ -\frac{(\nu_2 - \gamma_{i2})^2}{2(s_{i2}^2 + \sigma_2^2)} \right\}. \end{aligned}$$

Here s_{i2}^2 and γ_{i2} are the same functions as described in Section 3.1 and

$$b_i = \frac{r_{i21}^{r_{i21}} r_{i22}^{r_{i22}}}{(r_{i21}+r_{i22})^{r_{i21}+r_{i22}}} \left(\frac{2\pi(r_{i21}+r_{i22})}{r_{i21} r_{i22}} \right)^{1/2}.$$

The same method may be applied when there are no observed successes for subjects on treatment 2 ($n_{i21} = 0$) but no such extremes for subjects taking treatment 1 ($r_{i11} \notin \{0, n_{i1}\}$); in this case we obtain

$$I_i \approx a_i \frac{1}{[1+\exp(\nu_2)]^{n_{i2}}} \frac{1}{\sqrt{2\pi(s_{i1}^2 + \sigma_1^2)}} \exp\left\{ -\frac{(\nu_1 - \gamma_{i1})^2}{2(s_{i1}^2 + \sigma_1^2)} \right\}$$

with

$$a_i = \frac{r_{i11}^{r_{i11}} r_{i12}^{r_{i12}}}{(r_{i11}+r_{i12})^{r_{i11}+r_{i12}}} \left(\frac{2\pi(r_{i11}+r_{i12})}{r_{i11} r_{i12}} \right)^{1/2}.$$

Case 2: No observed failures in either treatment.

The same procedure as used for Case 1 may be followed when there are no observed failures in either drug or control ($r_{i11} = n_{i1}$ or $r_{i21} = n_{i2}$). For the case when $r_{i11} = n_{i1}$ and $r_{i21} \notin \{0, n_{i2}\}$ the integral is approximated by

$$I_i \approx b_i \frac{\exp(\nu_1)^{n_{i1}}}{(1+\exp(\nu_1))^{n_{i1}}} \frac{1}{\sqrt{2\pi(s_{i2}^2 + \sigma_2^2)}} \exp\left\{ -\frac{(\nu_2 - \gamma_{i2})^2}{2(s_{i2}^2 + \sigma_2^2)} \right\}.$$

For the case when $r_{i21} = n_{i2}$ and $r_{i11} \notin \{0, n_{i1}\}$ the integral is approximated by

$$I_i \approx a_i \frac{\exp(\nu_2)^{n_{i2}}}{(1+\exp(\nu_2))^{n_{i2}}} \frac{1}{\sqrt{2\pi(s_{i1}^2 + \sigma_1^2)}} \exp\left\{ -\frac{(\nu_1 - \gamma_{i1})^2}{2(s_{i1}^2 + \sigma_1^2)} \right\}.$$

Case 3: Observed extremes in both treatments.

For extreme observations on both treatments we may use the approximation (10) for both treatments. The integrand in I_i will then consist of only the bivariate normal distribution integrated over the whole range and will therefore be equal to one. For example, if both $r_{i11} = 0$ and $r_{i21} = 0$, the intergal I_i is approximated by

$$I_i \approx \frac{1}{(1+e^{\nu_1})^{n_{i1}}} \frac{1}{(1+e^{\nu_2})^{n_{i2}}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\boldsymbol{\mu}_i; \boldsymbol{\nu}, \mathbb{V}) d\boldsymbol{\mu}_i = \frac{1}{(1+e^{\nu_1})^{n_{i1}}} \frac{1}{(1+e^{\nu_2})^{n_{i2}}}.$$

5.4. *Retaining the quadratic form of the log-likelihood function.*

The use of the approximation (10) to the components of the log-likelihood function when $r_{ijq} = 0$ or $r_{ijq} = n_{ij}$ $j, q \in \{1, 2\}$ does not retain the quadratic-form of the log-likelihood. For example, the log-likelihood function for Case 1 where $r_{i11} = 0$ and $r_{i21} \notin \{0, n_{i2}\}$ is given by

$$\begin{aligned} \log I_i &\approx \log\left[b_i \frac{1}{(1+\exp(\nu_1))^{n_{i1}}} \frac{1}{\sqrt{2\pi(s_{i2}^2+\sigma_2^2)}} \exp\left\{ -\frac{(\nu_2-\gamma_{i2})^2}{2(s_{i2}^2+\sigma_2^2)} \right\} \right] \\ &= -n_{i1}\log\left(1+\exp(\nu_1)\right) - \frac{(\nu_2-\gamma_{i2})^2}{2(s_{i2}^2+\sigma_2^2)} + Const. \end{aligned}$$

Note that the approximation is a quadratic function in ν_2 but not in ν_1. The term involving ν_1 may be expanded using the Taylor series at $\nu_1 = 0$ to give

$$\log I_i \approx -n_{i1}\left(\log(2) + \frac{1}{2}\nu_1 + \frac{1}{8}\nu_1^2\right) - \frac{(\nu_2-\gamma_{i2})^2}{2(s_{i2}^2+\sigma_2^2)} + Const.$$

By using this approximation we maintain the quadratic form for the approximation to the log-likelihood function and hence the estimator retains the form:

$$\hat{\mu}_1 = \frac{K_3K_5 + K_2K_4}{K_1K_2 - K_3^2},$$

$$\hat{\mu}_2 = \frac{K_1K_5 + K_4K_3}{K_1K_2 - K_3^2}.$$

Adjustments, however, need to be made to $K_{1i}, K_{2i}, K_{3i}, K_{4i}$ and K_{5i} and these are given in Table 3. (The terms K_{li}^* of the table have to be added to respective K_{li}, $l = 1, \ldots, 5$.)

References

[1] AGRESTI, A. & HARTZEL, J. (2000) Strategies for comparing treatments on a binary response with multi-centre data. *Statistics in Medicine*, **19**, 1115–1139.

[2] DRAGALIN, V. & FEDOROV, V.(2004) Multi-center trials with binary response. *GSK BDS Technical Report,* 2004-03.

[3] FEDOROV, V. & JONES, B. (2005) Design and analysis of multicentre trials. *Statistical Methods in Medical Research*, **XX**. To appear.

Centre Estimator

Variance		Mean			$\sqrt{R}$ MSE		
Center	Treatment	(8)	SAS®	Difference	(8)	SAS®	Difference
0.5	1	0.1201	0.1217	-0.0016	0.3327	0.3447	-0.0120
	0.8	0.1305	0.1325	-0.0020	0.1879	0.1971	-0.0092
	0.6	0.0988	0.1002	-0.0014	0.3679	0.3808	-0.0129
	0.4	0.0951	0.0967	-0.0016	0.2564	0.2638	-0.0074
	0.2	0.0927	0.0942	-0.0015	0.3558	0.3663	-0.0105
0.25	1	0.0631	0.0642	-0.0011	0.2038	0.2063	-0.0025
	0.8	0.1224	0.1240	-0.0016	0.2199	0.2282	-0.0083
	0.6	0.0863	0.0877	-0.0014	0.1399	0.1431	-0.0032
	0.4	0.0960	0.0973	-0.0013	0.1671	0.1712	-0.0041
	0.2	0.0970	0.0981	-0.0011	0.0844	0.0866	-0.0022
0.1	1	0.1051	0.1065	-0.0014	0.0873	0.0897	-0.0024
	0.8	0.0941	0.0953	-0.0012	0.1286	0.1317	-0.0031
	0.6	0.1058	0.1073	-0.0015	0.0761	0.0786	-0.0025
	0.4	0.0978	0.0990	-0.0012	0.0915	0.0938	-0.0023
	0.2	0.0920	0.0931	-0.0011	0.0625	0.0634	-0.0009

Treatment Estimator

Variance		Mean			$\sqrt{R}$ MSE		
Center	Treatment	(8)	SAS®	Difference	(8)	SAS®	Difference
0.5	1	0.5173	0.5249	-0.0076	0.6384	0.6699	-0.0304
	0.8	0.5242	0.5332	-0.0090	0.3436	0.3698	-0.0262
	0.6	0.4912	0.4995	-0.0083	0.4971	0.5140	-0.0169
	0.4	0.4819	0.4893	-0.0074	0.2781	0.2783	-0.0002
	0.2	0.4941	0.5010	-0.0069	0.1938	0.1989	-0.0051
0.25	1	0.4396	0.4460	-0.0064	0.8085	0.8047	0.0038
	0.8	0.5361	0.5436	-0.0075	0.6833	0.7227	-0.0394
	0.6	0.4865	0.4931	-0.0066	0.3325	0.3379	-0.0054
	0.4	0.4946	0.5007	-0.0061	0.2999	0.3068	-0.0069
	0.2	0.4945	0.5006	-0.0061	0.1133	0.1152	-0.0019
0.1	1	0.5337	0.5410	-0.0073	0.7956	0.8346	-0.0390
	0.8	0.4763	0.4823	-0.0060	0.9187	0.9349	-0.0162
	0.6	0.5351	0.5419	-0.0068	0.5783	0.6079	-0.0296
	0.4	0.4837	0.4894	-0.0057	0.3670	0.3712	-0.0042
	0.2	0.4905	0.4964	-0.0059	0.1591	0.1608	-0.0017

TABLE 1. Mean values of estimators along with respective MSE's computed for the estimators (8) and using adaptive Gaussian quadratures in SAS® with 50 centres, 100 patients on each arm, $R = 1000$ runs, centre effect $\nu_0 = 0.1$, treatment effect $\delta = 0.5$, $\sigma_1 = \sigma_2$ (so that the covariance between centre and treatment effects is zero) and different centre and treatment variances.

Centre Estimator

Variance		Sample Variance			Mean Predicted Variance		
Center	Treatment	(8)	SAS®	Difference	(8)	SAS®	Difference
0.5	1	0.0101	0.0104	-0.0003	0.0105	0.0105	0.0000
	0.8	0.0050	0.0052	-0.0002	0.0105	0.0105	0.0000
	0.6	0.0116	0.0121	-0.0005	0.0105	0.0105	0.0000
	0.4	0.0081	0.0083	-0.0002	0.0105	0.0105	0.0000
	0.2	0.0112	0.0116	-0.0004	0.0105	0.0105	0.0000
0.25	1	0.0051	0.0052	-0.0001	0.0055	0.0055	0.0000
	0.8	0.0065	0.0066	-0.0001	0.0055	0.0055	0.0000
	0.6	0.0042	0.0044	-0.0002	0.0055	0.0055	0.0000
	0.4	0.0053	0.0054	-0.0001	0.0055	0.0054	0.0000
	0.2	0.0027	0.0027	0.0000	0.0054	0.0054	0.0000
0.1	1	0.0027	0.0028	-0.0001	0.0025	0.0024	0.0000
	0.8	0.0040	0.0041	-0.0001	0.0024	0.0024	0.0000
	0.6	0.0024	0.0024	0.0000	0.0024	0.0024	0.0000
	0.4	0.0029	0.0030	-0.0001	0.0024	0.0024	0.0000
	0.2	0.0019	0.0020	-0.0001	0.0024	0.0024	0.0000

Treatment Estimator

Variance		Sample Variance			Mean Predicted Variance		
Center	Treatment	(8)	SAS®	Difference	(8)	SAS®	Difference
0.5	1	0.0199	0.0205	-0.0006	0.0220	0.0220	0.0000
	0.8	0.0103	0.0106	-0.0003	0.0180	0.0180	0.0001
	0.6	0.0157	0.0163	-0.0006	0.0140	0.0139	0.0000
	0.4	0.0085	0.0087	-0.0002	0.0099	0.0099	0.0000
	0.2	0.0061	0.0063	-0.0002	0.0059	0.0059	0.0000
0.25	1	0.0219	0.0226	-0.0007	0.0219	0.0218	0.0000
	0.8	0.0203	0.0210	-0.0007	0.0179	0.0178	0.0000
	0.6	0.0103	0.0106	-0.0003	0.0138	0.0138	0.0000
	0.4	0.0095	0.0097	-0.0002	0.0098	0.0098	0.0000
	0.2	0.0036	0.0036	0.0000	0.0058	0.0058	0.0000
0.1	1	0.0240	0.0247	-0.0007	0.0218	0.0218	0.0000
	0.8	0.0285	0.0293	-0.0008	0.0178	0.0178	0.0000
	0.6	0.0171	0.0175	-0.0004	0.0138	0.0137	0.0000
	0.4	0.0114	0.0116	-0.0002	0.0097	0.0097	0.0000
	0.2	0.0049	0.0051	-0.0002	0.0057	0.0057	0.0000

TABLE 2. Sample variances and means of predicted variances for the estimators (8) and computed using adaptive Gaussian quadratures in SAS® with 50 centres, 100 patients on each arm, $R = 1000$ runs, centre effect $\nu_0 = 0.1$, treatment effect $\delta = 0.5$, $\sigma_1 = \sigma_2$ and different centre and treatment variances.

n_{i11}	n_{i21}	K^*_{1i}	K^*_{2i}	K^*_{3i}	K^*_{4i}	K^*_{5i}
0	$\notin \{0, n_{i2}\}$	$\frac{n_{i1}}{4}$	$\frac{1}{(s^2_{i2}+\sigma^2_2)}$	0	$-\frac{n_{i1}}{2}$	$\frac{\gamma_{i2}}{(s^2_{i2}+\sigma^2_2)}$
$\notin \{0, n_{i1}\}$	0	$\frac{1}{(s^2_{i1}+\sigma^2_1)}$	$\frac{n_{i2}}{4}$	0	$\frac{\gamma_{i1}}{(s^2_{i1}+\sigma^2_1)}$	$-\frac{n_{i2}}{2}$
n_{i1}	$\notin \{0, n_{i2}\}$	$\frac{n_{i1}}{4}$	$\frac{1}{(s^2_{i2}+\sigma^2_2)}$	0	$\frac{n_{i1}}{2}$	$\frac{\gamma_{i2}}{(s^2_{i2}+\sigma^2_2)}$
$\notin \{0, n_{i1}\}$	n_{i2}	$\frac{1}{(s^2_{i1}+\sigma^2_1)}$	$\frac{n_{i2}}{4}$	0	$\frac{\gamma_{i1}}{(s^2_{i1}+\sigma^2_1)}$	$\frac{n_{i2}}{2}$
0	0	$\frac{n_{i1}}{4}$	$\frac{n_{i2}}{4}$	0	$-\frac{n_{i1}}{2}$	$-\frac{n_{i2}}{2}$
n_{i1}	n_{i2}	$\frac{n_{i1}}{4}$	$\frac{n_{i2}}{4}$	0	$\frac{n_{i1}}{2}$	$\frac{n_{i2}}{2}$
0	n_{i2}	$\frac{n_{i1}}{4}$	$\frac{n_{i2}}{4}$	0	$-\frac{n_{i1}}{2}$	$\frac{n_{i2}}{2}$
n_{i1}	0	$\frac{n_{i1}}{4}$	$\frac{n_{i2}}{4}$	0	$\frac{n_{i1}}{2}$	$-\frac{n_{i2}}{2}$

TABLE 3. Zero Term Adjustments