TECHNICAL REPORT

Method-of-moments computations
for meta-analysis using summary (trial-level) data
when applied to binary surrogate and true endpoints

Supplement to: **A simple meta-analytic approach for using a binary surrogate endpoint to predict the effect of intervention on true endpoint**
Biostatistics 2005

Stuart G. Baker

Biometry Research Group, Division of Cancer Prevention, National Cancer Institute

We adapted the trial-level meta-analysis of surrogate endpoints (Gail et al, 2000, Buyse et al, 2000) to binary surrogate and true endpoints using a relatively simple method-of-moments approach for estimation. Let Z denote randomization group, S denote a surrogate endpoint with realizations of 0 or 1, and T denote a true endpoint with realizations 0 or 1. Also let $I = 0$ denote the application or validation trial and $I = i$ $= 1, 2,$ k index previous trials. We define

$\phi_{0zT} = pr(T = 1 | Z = z, I = 0) =$ the probability the true endpoint equals 1 in arm z of the application or validation trial,

$\phi_{0zS} = pr(S = 1 | Z = z, I = i) =$ the probability the surrogate endpoints equals 1 in arm z of the application or validation trial,

$\phi_{izS} = pr(S = 1 | Z = z, I = i), =$ the probability the true endpoint equals 1 in arm z of previous trial i,

$\phi_{izT} = pr(T = 1 | Z = z, I = i) =$ the probability the true endpoint equals 1 in arm z of previous trial i, The parameter of interest is the treatment effect in the new trial $\Delta_0 = \phi_{01T} - \phi_{00T}$. We construct a joint random effects model for the estimate of ϕ_{izS} and ϕ_{izT} based on data from previous trials $i = 1, 2, k$ and use the model along with the estimates in new trial of the probability of surrogate endpoint to estimate the treatment effect Δ_0 in the new trial. Let n_{izst} denote the number of subjects in group z of trial i who have surrogate outcome s and true outcome t. The estimates under a binomial model are $\widehat{\phi}_{izS} = n_{iz1+}/n_{iz++}$ and $\widehat{\phi}_{izT} = n_{iz+1}/n_{iz++}$. Let $\underset{\sim}{\phi_i} = (\Delta_i, \phi_{i0S}, \phi_{i1S})^T$, where $\Delta_i = \phi_{i1T} - \phi_{i0T}$. We assume that $\underset{\sim}{\widehat{\phi}_i} \sim N(\underset{\sim}{\phi_i}, V_{\text{sampling}(i)})$. Let $\theta_{izst} = pr(S = s, T = t | i, z)$. Applying the delta method

$$V_{\text{sampling}(i)} = \begin{pmatrix} \sigma^2_{i\Delta\Delta} & \sigma^2_{i\Delta S0} & \sigma^2_{i\Delta S1} \\ \sigma^2_{i\Delta S0} & \sigma^2_{iS0S0} & 0 \\ \sigma^2_{i\Delta S1} & 0 & \sigma^2_{iS1S1} \end{pmatrix} = A \cdot W_i \cdot A^T, \text{ where}$$

$$W_i = \begin{pmatrix} w_{i0SS} & w_{i0ST} & 0 & 0 \\ w_{i0ST} & w_{i0TT} & 0 & 0 \\ 0 & 0 & w_{i1SS} & w_{i1ST} \\ 0 & 0 & w_{i1ST} & w_{i1TT} \end{pmatrix}, A = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad (1)$$

$$w_{izSS} = \phi_{izS}(1 - \phi_{izS}) / n_{iz++},$$

$$w_{izTT} = \phi_{izT}(1 - \phi_{izT}) / n_{iz++},$$

$$w_{izST} = cov(\phi_{izS}, \phi_{izT}) = cov(\theta_{iz10} + \theta_{iz11}, \theta_{iz01} + \theta_{iz11})$$

$$= (1 \quad 0 \quad 1) \begin{pmatrix} \theta_{iz01}(1 - \theta_{iz01}) & -\theta_{iz01}\theta_{iz10} & -\theta_{iz01}\theta_{iz11} \\ -\theta_{iz01}\theta_{iz10} & \theta_{iz10}(1 - \theta_{iz10}) & -\theta_{iz10}\theta_{iz11} \\ -\theta_{iz01}\theta_{iz11} & -\theta_{iz10}\theta_{iz11} & \theta_{iz11}(1 - \theta_{iz11}) \end{pmatrix} / n_{iz++} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$= \{ \theta_{iz11} (1 - \phi_{izS}) - \theta_{iz01} \phi_{izS} \} / n_{iz++},$$

For estimation, $\widehat{\theta}_{iz11} = n_{iz11}/n_{iz++}$ and $\widehat{\theta}_{iz01} = n_{iz01}/n_{iz++}$.

Following the general approach of Gail et al (2000) and Buyse et al (2000), to allow for extra variability, we superimpose a random effects models, $\underset{\sim}{\phi_i} \sim N(\underset{\sim}{\phi}, V_{\text{random}})$, where

$$V_{\text{random}} = \begin{pmatrix} d_{\Delta\Delta} & d_{\Delta S0} & d_{\Delta S1} \\ d_{\Delta S0} & d_{S0S0} & d_{S0S1} \\ d_{\Delta S1} & d_{S0S1} & d_{S1S1} \end{pmatrix}. \tag{2}$$

Let $\underset{\sim}{\omega_0} = (\Delta_0, \widehat{\phi}_{00S}, \widehat{\phi}_{01S})$, where $\Delta_0 = \phi_{01T} - \phi_{00T}$ is the parameter of interest. Based on $(B1)$ and $(B2)$, we assume a joint normal distribution for $\underset{\sim}{\omega_0}$ with mean $(E(\Delta_0), E(\phi_{00S}), E(\phi_{01S}))$ and variance

$$\begin{pmatrix} d_{\Delta\Delta} & d_{\Delta S0} & d_{\Delta S1} \\ d_{\Delta S0} & d_{S0S0}+\sigma^2_{0S0S0} & d_{S0\text{S}1} \\ d_{\Delta S1} & d_{S0\text{S}1} & d_{S1S1}+\sigma^2_{0S1S1} \end{pmatrix} \tag{3}$$

Based on the formula for conditional joint normal distribution (e.g. Morrison, 1980, p. 97),

$$E(\Delta_0 \mid \widehat{\phi}_{00S}, \widehat{\phi}_{01S}) = E(\Delta_0) + (d_{\Delta\text{S}0}, d_{\Delta S1}) \begin{pmatrix} d_{S0S0}+\sigma^2_{0S0S0} & d_{S0\text{S}1} \\ d_{S0\text{S}1} & d_{S1S1}+\sigma^2_{0S1S1} \end{pmatrix}^{-1} \left(\begin{pmatrix} \widehat{\phi}_{00S} \\ \widehat{\phi}_{01S} \end{pmatrix} - \begin{pmatrix} E(\phi_{00S}) \\ E(\phi_{01S}) \end{pmatrix} \right) \tag{4}$$

The formula in Buyse et al (2000) is (4) with $\sigma^2_{0S0S0} = \sigma^2_{0S1S1} = 0$, which ignores the variability in the estimates from the application or validation trial, although, in practice,we found this made little difference.

We plug the following estimates into (4). We estimate $E(\Delta_0)$ by $\Sigma_i \, \widehat{\Delta}_i \, / k$, and we estimate $E(\widehat{\phi}_{0zS})$ by $\Sigma_i \, \widehat{\phi}_{izS} \, / k$. We estimate the components of $\widehat{V}_{\text{random}}$ using the following method-of-moments approach. Let $\underset{\sim}{\overline{\phi}} = \Sigma i \, \underset{\sim}{\widehat{\phi}_i} \, / \, k$. Without imposing any model the estimated variance of $\overline{\underset{\sim}{\phi}}$ is $\widehat{V}_{\text{total}} = (\widehat{\underset{\sim}{\phi}}_i - \overline{\underset{\sim}{\phi}})^T \cdot (\widehat{\underset{\sim}{\phi}}_i - \overline{\underset{\sim}{\phi}}) /(k\,(k-1))$. The sampling variance for $\overline{\underset{\sim}{\phi}}$ is $\widehat{V}_{\text{sampling}} = \Sigma_i \widehat{V}_{\text{sampling}(i)} / k^2$, where $\widehat{V}_{\text{sampling}(i)}$ is simply $V_{\text{sampling}(i)}$ with the estimates substituted for the parameters. The method of moments estimate of $\widehat{V}_{\text{random}}$ is $\widehat{V}_{\text{random}} = \widehat{V}_{\text{total}} - \widehat{V}_{\text{sampling}}$. However for $\widehat{V}_{\text{random}}$ to be a valid estimated covariance matrix, it must be positive definite. If $\widehat{V}_{\text{random}}$ is not positive definite, we compute an adjusted covariance matrix V^*_{random} as follows. We first write $\widehat{V}_{\text{random}} = M\, Diagonal\,(\underset{\sim}{\lambda})\; M$, where $\underset{\sim}{\lambda} = (\lambda_1,\, \lambda_2\,, \ldots.)$ are the eigenvalues, $Diagonal$ indicates a diagonal matrix, and M is the normalized eigenmatrix. We then replace eigenvalues that are less than or equal to zero with a small positive value to obtain a new vector of eigenvalues $\underset{\sim}{\lambda^*}$. We then computed the adjusted positive definite covariance matrix $V^*_{\text{random}} = M\, Diagonal\,(\underset{\sim}{\lambda^*})\; M$.

Additional Reference

Morrison DF. (1976). *Multivariate Statistical Methods.* New York: McGraw-Hill Book Company.