Appendix

A.1 Variance formulas

A.1.1 Proof of formulas (3.2) and (3.3)

From equation (2.1) we have $\mathbf{\Sigma}_{\mathrm{B}} = \mathbb{E}^{-1}\left(\mathbf{X}'_i \mathbf{\Sigma}_i^{-1} \mathbf{X}_i\right)$. Now,

$$\mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i =$$

$$\begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ t_{i0} & \cdots & t_{i0}+sj & \cdots & t_{i0}+sr \\ k_i & \cdots & k_i & \cdots & k_i \end{pmatrix} \begin{pmatrix} \nu_{00} & \cdots & \nu_{0r} \\ \vdots & \ddots & \vdots \\ \nu_{r0} & \cdots & v_{rr} \end{pmatrix} \begin{pmatrix} 1 & t_{i0} & k_i \\ \vdots & \vdots & \vdots \\ 1 & t_{i0}+sj & k_i \\ \vdots & \vdots & \vdots \\ 1 & t_{i0}+sr & k_i \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & & \\ \left\{ t_{i0}\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} \right\} & \left\{ t_{i0}^2 \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s t_{i0}\sum_{j=0}^{r}\sum_{j'=0}^{r} (j+j')v_{jj'} + s^2 \sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} \right\} & \\ k_i \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & k_i \left(t_{i0}\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} \right) & k_i^2 \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} \end{pmatrix}$$

Using $\bar{t}_0 = \mathbb{E}(t_0)$, $p_e = \mathbb{E}(k) = \mathbb{E}(k^2)$ and

$$\rho_{\mathrm{e},t_0} = \frac{\mathbb{E}(kt_0) - p_e \bar{t}_0}{\sqrt{p_e(1-p_e)}\sqrt{V(t_0)}},$$

and assuming without loss of generality that $\bar{t}_0 = 0$ (this can be achieved by centering the initial time), which implies $\mathbb{E}(t_0^2) = V(t_0)$, we have that

$$\mathbb{E}\left(\mathbf{X}'_i \mathbf{\Sigma}_i^{-1} \mathbf{X}_i\right) =$$

$$\begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & & \\ s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & V(t_0)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s^2 \sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} & \\ p_e \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & \left(\rho_{\mathrm{e},t_0}\sqrt{p_e(1-p_e)}\sqrt{V(t_0)}\right)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s p_e \sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & p_e \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} \end{pmatrix}$$

We are interested in the [3,3] component of the inverse of this matrix, which is

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \mathbf{c}'\left(\mathbb{E}\left(\mathbf{X}'_i\boldsymbol{\Sigma}_i\mathbf{X}_i\right)\right)^{-1}\mathbf{c} = \frac{s^2\det(\mathbf{A}) + \left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)^2 V(t_0)}{p_e(1-p_e)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)\left[s^2\det(\mathbf{A}) + \left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)^2\left(1-\rho_{e,t_0}^2\right)V(t_0)\right]}$$

If either $V(t_0)$ or ρ_{e,t_0} are zero then

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{1}{p_e(1-p_e)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)}.$$

If we follow Lachin's approach (Lachin, 2000), instead of using the asymptotic variance use the variance of $\hat{\mathbf{B}}$ conditional on the covariates, which is

$$\left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right)^{-1},$$

and redefine $\boldsymbol{\Sigma}_{\mathrm{B}}$ as

$$\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right)^{-1}$$

so that the test statistic is still

$$T = \frac{\sqrt{N}\,\mathbf{c}'\hat{\mathrm{B}}}{\sqrt{\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}}}.$$

Then, we would take the expected value of the non-centrality parameter under the alternative hypothesis over the distribution of $\mathbf{X}_i$, i.e. we would compute $\mathbb{E}\left[T^2|H_1\right]$. If we assume that everyone is observed at the same set of time points, then the only random covariate is exposure. Thus,

$$\frac{1}{N}\sum_i\left(\mathbf{X}_i\boldsymbol{\Sigma}^{-1}\mathbf{X}'_i\right) =$$

$$\begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & & \\ s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & s^2\sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} & \\ \left(\frac{\sum_i k_i}{N}\right)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & s\left(\frac{\sum_i k_i}{N}\right)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'}\right) & \left(\frac{\sum_i k_i^2}{N}\right)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} \end{pmatrix},$$

and the [3,3] component of the inverse is

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \left[\left(\frac{\sum_i k_i}{N} \right) \left(1 - \frac{\sum_i k_i}{N} \right) \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \right]^{-1} .$$

Then,

$$T^2 = \frac{\hat{\beta}_2^2}{Var(\hat{\beta}_2)} = N\hat{\beta}_2^2 \left(\frac{\sum_i k_i}{N} \right) \left(1 - \frac{\sum_i k_i}{N} \right) \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right)$$

and

$$\mathbb{E}\left[T^2|H_1\right] = \mathbb{E}\left[N\beta_2^2 \left(\frac{\sum_i k_i}{N} \right) \left(1 - \frac{\sum_i k_i}{N} \right) \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \right]$$

where the expected value is taken over the distribution of k_i, so

$$\mathbb{E}\left[T^2|H_1\right] = N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \mathbb{E}\left[\left(\frac{\sum_i k_i}{N} \right) \left(1 - \frac{\sum_i k_i}{N} \right) \right] .$$

Noticing that $Z = \sum_i k_i$ is a Binomial variable we can work out the expected value,

$$\begin{aligned}
\mathbb{E}\left[T^2|H_1\right] &= N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \mathbb{E}\left[\left(\frac{Z}{N} \right) \left(1 - \frac{Z}{N} \right) \right] \\
&= N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \left(\frac{\mathbb{E}(Z)}{N} - \frac{\mathbb{E}(Z^2)}{N^2} \right) \\
&= N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \left(p_e - \frac{Np_e(1-p_e) + N^2 p_e^2}{N^2} \right) \\
&= N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) \left(\frac{Np_e - p_e + p_e^2 - Np_e^2}{N} \right) \\
&= (N-1)\,\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) p_e\,(1-p_e) .
\end{aligned}$$

The non-centrality parameter with the approach we followed in the paper is

$$N\beta_2^2 \left(\sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right) p_e\,(1-p_e) ,$$

so there is only a $\left(1 - \frac{1}{N}\right)$ correction compared with the one obtained with Lachin's method.

A.1.2 Proof of formula (3.4)

Following model (2.6), and our derivations on Appendix A.1.1, we now have that

$$\mathbf{X}'_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i = \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ t_{i0} & \cdots & t_{i0}+sj & \cdots & t_{i0}+sr \\ k_i & \cdots & k_i & \cdots & k_i \\ k_i t_{i0} & \cdots & k_i t_{i0} + k_i sj & \cdots & k_i t_{i0} + k_i sr \end{pmatrix} \begin{pmatrix} v_{00} & \cdots & v_{0r} \\ \vdots & \ddots & \vdots \\ v_{r0} & \cdots & v_{rr} \end{pmatrix} \begin{pmatrix} 1 & t_{i0} & k_i & k_i t_{i0} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{i0}+sj & k_i & k_i t_{i0} + k_i sj \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{i0}+sr & k_i & k_i t_{i0} + k_i sr \end{pmatrix}$$

and using the results in Appendix A.1.1 we only need to derive the components in the last row. We can derive $\mathbb{E}\left(\mathbf{X}'_i \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i\right)$, in which the [4,1] component is equivalent to the [3,2] and therefore it takes the value

$$\left(\rho_{e,t_0} \sqrt{p_e(1-p_e)} \sqrt{V\left(t_0\right)}\right) \sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} + s p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j v_{jj'}.$$

The [4,2] component is

$$\mathbb{E}\left(kt_0^2\right) \sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} + 2s \left(\rho_{e,t_0} \sqrt{p_e(1-p_e)} \sqrt{V\left(t_0\right)}\right) \sum_{j=0}^{r} \sum_{j'=0}^{r} j v_{jj'} + s^2 p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'},$$

the [4,3] component is

$$\left(\rho_{e,t_0} \sqrt{p_e(1-p_e)} \sqrt{V\left(t_0\right)}\right) \sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} + s p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j v_{jj'},$$

and the [4,4] component is the same as the [4,2] component. An expression for $\mathbb{E}\left(kt_0^2\right) = p_e \mathbb{E}\left(t_{0,k=1}^2\right)$ in terms of the known parameters is needed. Since we assumed that $\bar{t}_0 = 0$, then $V\left(t_0\right) = \mathbb{E}\left(t_0^2\right) = (1-p_e)\mathbb{E}\left(t_{0,k=0}^2\right) + p_e \mathbb{E}\left(t_{0,k=1}^2\right)$, which implies

$$\mathbb{E}\left(t_{0,k=0}^2\right) = \frac{V\left(t_0\right) - p_e \mathbb{E}\left(t_{0,k=1}^2\right)}{1-p_e} \tag{A.1}$$

We have from Appendix A.1.1 that

$$\mathbb{E}\left(kt_0\right) = p_e \bar{t}_{0,k=1} = \rho_{\mathrm{e},t_0}\sqrt{p_e(1-p_e)}\sqrt{V\left(t_0\right)},$$

therefore

$$\bar{t}_{0,k=1} = \rho_{\mathrm{e},t_0}\sqrt{\left(\frac{1-p_e)}{p_e}\right.}\sqrt{V\left(t_0\right)}$$

and it can be deduced that

$$\bar{t}_{0,k=0} = -\rho_{\mathrm{e},t_0}\sqrt{\left(\frac{p_e}{1-p_e)}\right.}\sqrt{V\left(t_0\right)}.$$

Then,

$$\left(\bar{t}_{0,k=1}\right)^2 = \rho^2_{\mathrm{e},t_0}\frac{(1-p_e)}{p_e}V\left(t_0\right)$$

and

$$\left(\bar{t}_{0,k=0}\right)^2 = \rho^2_{\mathrm{e},t_0}\frac{p_e}{(1-p_e)}V\left(t_0\right).$$

We assume that the variance of t_0 is the same in exposed and unexposed, i.e. $V\left(t_{0,k=0}\right) = V\left(t_{0,k=1}\right)$. It follows that

$$V\left(t_{0,k=0}\right) = V\left(t_{0,k=1}\right) \Leftrightarrow \mathbb{E}\left(t^2_{0,k=0}\right) - \left(\bar{t}_{0,k=0}\right)^2 = \mathbb{E}\left(t^2_{0,k=1}\right) - \left(\bar{t}_{0,k=1}\right)^2.$$

Plugging in expression (A.1) we obtain that

$$\mathbb{E}\left(t^2_{0,k=1}\right) = V\left(t_0\right) + \rho^2_{\mathrm{e},t_0}\left(\frac{1-2p_e}{p_e}\right)V\left(t_0\right).$$

Therefore,

$$\mathbb{E}\left(kt_0^2\right) = p_e\mathbb{E}\left(t^2_{0,k=1}\right) = V\left(t_0\right)\left[p_e + \rho^2_{\mathrm{e},t_0}(1-2p_e)\right].$$

Now, plugging in this last expression in the formula for $\mathbb{E}\left(\mathbf{X}'_i\mathbf{\Sigma}_i^{-1}\mathbf{X}_i\right)$, and inverting the matrix, it can be derived that its [4,4] component is

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \mathbf{c}'\left(\mathbb{E}\left(\mathbf{X}'_i\mathbf{\Sigma}_i^{-1}\mathbf{X}_i\right)\right)^{-1}\mathbf{c} =$$

$$\frac{\sum_{j=0}^{r}\sum_{j'=0}^{r}\upsilon_{jj'}}{p_e(1-p_e)\left[s^2\det(\mathbf{A}) + \left(1-\rho^2_{\mathrm{e},t_0}\right)V(t_0)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r}\upsilon_{jj'}\right)^2\right]}.$$

If $V(t_0) = 0$, then

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum_{j=0}^{r}\sum_{j'=0}^{r}\upsilon_{jj'}}{p_e(1-p_e)s^2\det(\mathbf{A})},$$

and if $\rho_{\mathrm{e},t_0} = 0$ then

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum_{j=0}^{r}\sum_{j'=0}^{r}\upsilon_{jj'}}{p_e(1-p_e)\left[s^2\det(\mathbf{A}) + V(t_0)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r}\upsilon_{jj'}\right)^2\right]}.$$

If we can assume that t_0 and exposure are independent, then the formula we derived for the case $\rho_{\mathrm{e},t_0} = 0$ also applies to model (2.7), which assumes a general form for the relationship between response and time in the unexposed but requires that a main effect of time is in the model, we can rewrite the model as

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \gamma_0 + \gamma_1 t_{ij} + \alpha_1 f_1\left(t_{ij}\right) + \cdots + \alpha_q f_q\left(t_{ij}\right) + \gamma_2 k_i + \gamma_3\left(t_{ij}\times k_i\right),$$

where $f_u\left(t_{ij}\right),\ u = 1,\ldots,U$ are arbitrary functions of time. Since the $[m,q]$ term of the matrix $\mathbb{E}\left(\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)$ can be written as $\sum_{j,j'}\upsilon_{jj'}\mathbb{E}\left(x_{ijm}x_{ij'q}\right)$, where x_{kijm} is the value of the mth covariate for subject i from group k at time t_j, and exposure and time are independent, which implies

$$\mathbb{E}\left(k_i f_u\left(t_{ij'}\right)\right) = \mathbb{E}\left(k_i\right)\mathbb{E}\left(f_u\left(t_{ij'}\right)\right) = p_e\mathbb{E}\left(f_u\left(t_{ij'}\right)\right) \quad \forall u,$$

we have

$$
\mathbb{E}\left(\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)=
\begin{pmatrix}
\sum_{j,j'} v_{jj'} & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij'}\right)\right) & \cdots & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'} & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) \\
\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}t_{ij'}\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_1\left(t_{ij'}\right)\right) & \cdots & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}t_{ij'}\right) \\
\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij}\right)\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_1\left(t_{ij}\right)\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij}\right)f_1\left(t_{ij'}\right)\right) & \cdots & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij}\right)f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_1\left(t_{ij'}\right)\right) \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_V\left(t_{ij'}\right)\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_V\left(t_{ij'}\right)\right) & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij}\right)f_V\left(t_{ij'}\right)\right) & \cdots & \sum_{j,j'} v_{jj'}\mathbb{E}\left(f_V\left(t_{ij}\right)f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_V\left(t_{ij'}\right)\right) \\
p_e\sum_{j,j'} v_{jj'} & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_1\left(t_{ij}\right)\right) & \cdots & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(f_V\left(t_{ij}\right)\right) & p_e\sum_{j,j'} v_{jj'} & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) \\
p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}t_{ij'}\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_1\left(t_{ij}\right)\right) & \cdots & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}f_V\left(t_{ij'}\right)\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}\right) & p_e\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}t_{ij'}\right)
\end{pmatrix}
$$

$$
=\mathbf{M}=\begin{pmatrix}
\underset{(2\times(V+2))}{\mathbf{M}_1} & \\
\underset{(V\times(V+2))}{\mathbf{M}_2} & \underset{((V+2)\times 2)}{p_e\mathbf{M}'_1} \\
\underset{(2\times(V+2))}{p_e\mathbf{M}_1} & \underset{(2\times 2)}{p_e\mathbf{M}_4}
\end{pmatrix}
$$

Define the matrix

$$
\mathbf{Q}=\begin{pmatrix}
1 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & 1 & \ddots & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & \vdots \\
0 & \cdots & 0 & 1 & \ddots & \vdots \\
-p_e & 0 & \cdots & 0 & 1 & 0 \\
0 & -p_e & \cdots & 0 & 0 & 1
\end{pmatrix}
=\begin{pmatrix}
\underset{((V+2)\times(V+2))}{\mathbf{I}} & \underset{((V+2)\times 2)}{\mathbf{0}} \\
\underset{(2\times(V+2))}{\mathbf{Q}_1} & \underset{(2\times 2)}{\mathbf{I}}
\end{pmatrix},
$$

such that

$$\mathbf{QMQ}' = \mathbf{B} =$$

$$\begin{pmatrix} \mathbf{B}_1 & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ 0 \ \cdots \ 0 & p_e(1-p_e)\sum_{jj'} v_{jj'} & p_e(1-p_e)\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij'}\right) \\ 0 \ \cdots \ 0 & p_e(1-p_e)\sum_{jj'} v_{jj'}\mathbb{E}\left(t_{ij}\right) & p_e(1-p_e)\sum_{j,j'} v_{jj'}\mathbb{E}\left(t_{ij}t_{ij'}\right) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix}$$

Since $\mathbb{E}\left(t_{ij'}\right) = \mathbb{E}\left(t_0\right) + sj'$ and $\mathbb{E}\left(t_{ij}t_{ij'}\right) = \mathbb{E}\left(t_0^2\right) + s(j+j')\mathbb{E}\left(t_0\right) + s^2jj'$, and assuming without loss of generality that $\mathbb{E}\left(t_0\right) = 0$ and therefore $\mathbb{E}\left(t_0^2\right) = V(t_0)$, we have that

$$\mathbf{B}_2 = p_e(1-p_e)\begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & s\sum_{j=0}^{r}\sum_{j'=0}^{r} jv_{jj'} \\ s\sum_{j=0}^{r}\sum_{j'=0}^{r} jv_{jj'} & V\left(t_0\right)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s^2\sum_{j=0}^{r}\sum_{j'=0}^{r} jj'v_{jj'} \end{pmatrix}.$$

We are interested in the [V+4,V+4] component of $\mathbf{M}^{-1}$, which corresponds to $NVar\left(\hat{\gamma}_3\right)$. Now,

$$\mathbf{M}^{-1} = \mathbf{Q}'\mathrm{B}^{-1}\mathbf{Q} = \begin{pmatrix} \mathbf{B}_1^{-1} + \mathbf{Q}_1'\mathbf{B}_2^{-1}\mathbf{Q}_1 & \mathbf{Q}_1'\mathbf{B}_2^{-1} \\ \mathbf{B}_2^{-1}\mathbf{Q}_1 & \mathbf{B}_2^{-1} \end{pmatrix},$$

and

$$\mathbf{B}_2^{-1} = \frac{1}{p_e(1-p_e)\left[V(t_0)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)^2 + s^2\det(\mathbf{A})\right]}$$

$$\begin{pmatrix} V\left(t_0\right)\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} + s^2\sum_{j=0}^{r}\sum_{j'=0}^{r} jj'v_{jj'} & -s\sum_{j=0}^{r}\sum_{j'=0}^{r} jv_{jj'} \\ -s\sum_{j=0}^{r}\sum_{j'=0}^{r} jv_{jj'} & \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} \end{pmatrix}.$$

Thus, the [V+4,V+4] component of $\mathbf{M}^{-1}$ is

$$\frac{\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}}{p_e(1-p_e)\left[V(t_0)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)^2 + s^2\det(\mathbf{A})\right]}.$$

If we follow Lachin's approach (Lachin, 2000), instead of using the asymptotic variance use the variance of $\hat{\mathbf{B}}$ conditional on the covariates, which is

$$\left(\sum_{i=1}^{N} \mathbf{X}'_i \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i\right)^{-1},$$

and redefine $\boldsymbol{\Sigma}_{\mathrm{B}}$ as

$$\left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{X}'_i \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i\right)^{-1}$$

so that the test statistic is still

$$T = \frac{\sqrt{N}\,\mathbf{c}'\hat{\mathrm{B}}}{\sqrt{\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}}}.$$

Then, we would take the expected value of the non-centrality parameter under the alternative hypothesis over the distribution of $\mathbf{X}_i$, i.e. we would compute $\mathbb{E}\left[T^2|H_1\right]$. If we assume that everyone is observed at the same set of time points ($V(t_0) = 0$), then the only random covariate is exposure and we have

$$\frac{1}{N}\sum_i \mathbf{X}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}'_i =$$

$$\begin{pmatrix} \sum_{j,j'} v_{jj'} & & & \\ s\sum_{j,j'} j v_{jj'} & s^2 \sum_{j,j'} j j' v_{jj'} & & \\ \left(\frac{\sum_i k_i}{N}\right)\sum_{j,j'} v_{jj'} & s\left(\frac{\sum_i k_i}{N}\right)\sum_{j,j'} j v_{jj'} & \left(\frac{\sum_i k_i^2}{N}\right)\sum_{j,j'} v_{jj'} & \\ s\left(\frac{\sum_i k_i}{N}\right)\sum_{j,j'} j v_{jj'} & \left(\frac{s^2\sum_i k_i}{N}\right)\sum_{j,j'} j j' v_{jj'} & \left(\frac{s\sum_i k_i}{N}\right)\sum_{j,j'} j v_{jj'} & \left(\frac{s^2\sum_i k_i}{N}\right)\sum_{j,j'} j j' v_{jj'} \end{pmatrix}$$

and the [4,4] component of the inverse is

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}}{s^2 \det(\mathbf{A})\left(\frac{\sum_i k_i}{N}\right)\left(1 - \frac{\sum_i k_i}{N}\right)}.$$

Following the same steps as in Appendix A.1.1 we can derive that

$$\mathbb{E}\left[T^2|H_1\right] = \frac{(N-1)\,\gamma_3^2 s^2 \det(\mathbf{A}) p_e\,(1-p_e)}{\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}}.$$

The non-centrality parameter with the approach we followed in the paper is

$$\frac{N\gamma_3^2 s^2 \det(\mathbf{A}) p_e (1-p_e)}{\sum_{j=0}^{r}\left(\sum_{j'=0}^{r} v_{jj'}\right)},$$

so there is only a $\left(1-\frac{1}{N}\right)$ correction compared with the one obtained with Lachin's method.

A.1.3 Proof that $N\,Var\left(\hat{\eta}_5\right) = \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum_{j=0}^{r}\left(\sum_{j'=0}^{r} v_{jj'}\right)}{p_e(1-p_e)s^2\det(\mathbf{A})}$ under model (2.9).

From model (2.9), we have

$$\mathbf{X'}_i\mathbf{\Sigma}^{-1}\mathbf{X}_i = \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ t_{i0} & \cdots & t_{i0} & \cdots & t_{i0} \\ 0 & \cdots & sj & \cdots & sr \\ k_i & \cdots & k_i & \cdots & k_i \\ k_i t_{i0} & \cdots & k_i t_{i0} & \cdots & k_i t_{i0} \\ 0 & \cdots & k_i sj & \cdots & k_i sr \end{pmatrix} \begin{pmatrix} v_{00} & \cdots & v_{0r} \\ \vdots & \ddots & \vdots \\ v_{r0} & \cdots & v_{rr} \end{pmatrix} \begin{pmatrix} 1 & t_{i0} & 0 & k_i & k_i t_{i0} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{i0} & sj & k_i & k_i t_{i0} & k_i sj \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{i0} & sr & k_i & k_i t_{i0} & k_i sr \end{pmatrix}$$

and we can deduce using the following results derived in appendices 1.1 and 1.2, i.e.

$$\mathbb{E}\left(t_0\right) = \bar{t}_0 = 0, \mathbb{E}\left(t_0^2\right) = V\left(t_0\right), \mathbb{E}\left(k\right) = \mathbb{E}\left(k^2\right) = p_e,$$

$$\mathbb{E}\left(kt_0\right) = \rho_{\mathrm{e},t_0}\sqrt{p_e(1-p_e)}\sqrt{V\left(t_0\right)} = \overline{kt},$$

$$\mathbb{E}\left(kt_0^2\right) = V\left(t_0\right)\left[p_e + \rho_{\mathrm{e},t_0}^2(1-2p_e)\right] = \overline{kt^2},$$

that

$$\mathbb{E}\left(\mathbf{X}'_i\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right) =$$
$$\begin{pmatrix}
\sum\limits_{j,j'} v_{jj'} & & & & & \\
0 & V(t_0)\sum\limits_{j,j} v_{jj'} & & & & \\
s\sum\limits_{j,j'} jv_{jj'} & 0 & s^2\sum\limits_{j,j'} jj'v_{jj'} & & & \\
p_e\sum\limits_{j,j'} v_{jj'} & \overline{kt}\sum\limits_{j,j'} v_{jj'} & sp_e\sum\limits_{j,j'} jv_{jj'} & p_e\sum\limits_{j,j'} v_{jj'} & & \\
\overline{kt}\sum\limits_{j,j'} v_{jj'} & \overline{kt^2}\sum\limits_{j,j'} v_{jj'} & s\overline{kt}\sum\limits_{j,j'} jv_{jj'} & \overline{kt}\sum\limits_{j,j'} v_{jj'} & \overline{kt^2}\sum\limits_{j,j'} v_{jj'} & \\
sp_e\sum\limits_{j,j'} jv_{jj'} & s\overline{kt}\sum\limits_{j,j'} jv_{jj'} & s^2p_e\sum\limits_{j,j'} jj'v_{jj'} & sp_e\sum\limits_{j,j'} jv_{jj'} & s\overline{kt}\sum\limits_{j,j'} jv_{jj'} & s^2p_e\sum\limits_{j,j'} jj'v_{jj'}
\end{pmatrix}$$

The [6,6] component of the inverse of this matrix is

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum\limits_{j=0}^{r}\sum\limits_{j'=0}^{r} v_{jj'}}{p_e(1-p_e)s^2\det(\mathbf{A})},$$

as we derived in Appendix A.1.2 for the LDD case with $V(t_0) = 0$.

A.1.4 Proof that $s\hat{\lambda}_1 = \hat{\eta}_5$ and $s^2 Var\left(\hat{\lambda}_1\right) = Var\left(\hat{\eta}_5\right)$ from models (2.9) and (2.10)

The GLS estimator has the expression

$$\hat{\mathbf{B}} = \left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{Y}_i\right)$$

where $\mathbf{X}_i$ is the matrix of covariates for participant i. To derive $\hat{\eta}_5$ from model (2.9) we only need the sixth row of

$$\left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)^{-1},$$

which we denote

$$\left[\left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)^{-1}\right]_{[6]},$$

and then

$$\hat{\eta}_5 = \left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \right]_{[6]} \left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i \right)$$

which we rewrite as

$$\hat{\eta}_5 = \left(\sum_{i=1}^{N} \left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \right]_{[6]} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i \right).$$

Then, by calling

$$\mathbf{c}_\eta = \left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \right]_{[6]} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1}$$

we have

$$\hat{\eta}_5 = \left(\sum_{i=1}^{N} \mathbf{c}_\eta \mathbf{Y}_i \right)$$

In Appendix A.1.3 we derived and expression for

$$\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)$$

and from that we can derive

$$\left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \right]_{[6]} = \frac{1}{\det(\mathbf{A}) p_e (1-p_e) s} \left(p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j v_{jj'}, \quad 0, \quad \tfrac{-p_e}{s} \sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'}, \quad -\sum_{j=0}^{r} \sum_{j'=0}^{r} j v_{jj'}, \quad 0, \quad \tfrac{1}{s} \sum_{j=0}^{r} \sum_{j'=0}^{r} v_{jj'} \right)$$

For convenience, some terms can be rewritten in vector form. We define $\mathbf{1}$ as a $(r+1) \times 1$ vector of ones, and $\mathbf{t}$ as a $(r+1) \times 1$ matrix such that $\mathbf{t}' = (0, 1, 2, \ldots, r)$, and then

$$\left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \right]_{[6]} = \frac{1}{\det(\mathbf{A}) p_e (1-p_e) s} \left(p_e \mathbf{t}' \boldsymbol{\Sigma}^{-1} \mathbf{1}, \quad 0, \quad \tfrac{-p_e}{s} \mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}, \quad -\mathbf{t}' \boldsymbol{\Sigma}^{-1} \mathbf{1}, \quad 0, \quad \tfrac{1}{s} \mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1} \right)$$

We can also derive

$$\mathbf{X}'_i\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ t_{i0} & \cdots & t_{i0} & \cdots & t_{i0} \\ 0 & \cdots & sj & \cdots & sr \\ k_i & \cdots & k_i & \cdots & k_i \\ k_i t_{i0} & \cdots & k_i t_{i0} & \cdots & k_i t_{i0} \\ 0 & \cdots & k_i sj & \cdots & k_i sr \end{pmatrix} \begin{pmatrix} v_{00} & \cdots & v_{0r} \\ \vdots & \ddots & \vdots \\ v_{r0} & \cdots & v_{rr} \end{pmatrix} = \begin{pmatrix} \mathbf{1}'\boldsymbol{\Sigma}^{-1} \\ t_{i0}\mathbf{1}'\boldsymbol{\Sigma}^{-1} \\ s\mathbf{t}'\boldsymbol{\Sigma}^{-1} \\ k_i\mathbf{1}'\boldsymbol{\Sigma}^{-1} \\ k_i t_{i0}\mathbf{1}'\boldsymbol{\Sigma}^{-1} \\ sk_i\mathbf{t}'\boldsymbol{\Sigma}^{-1} \end{pmatrix}.$$

Then,

$$\mathbf{c}_\eta = \left[\left(\sum_{i=1}^{N} \mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i \right)^{-1} \right]_{[6]} \mathbf{X}'_i\boldsymbol{\Sigma}^{-1} = \frac{1}{\det(\mathbf{A})p_e(1-p_e)s}$$
$$\left(p_e\mathbf{t}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1} - p_e\mathbf{t}'\boldsymbol{\Sigma}^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1} - k_i\mathbf{t}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1} + k_i\mathbf{t}'\boldsymbol{\Sigma}^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right)$$
$$= \frac{(-p_e+k_i)\left(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right)}{\det(\mathbf{A})p_e(1-p_e)s}\mathbf{t}'\left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\right).$$

Now let us move to model (2.10). Define the $r \times (r+1)$ matrix

$$\boldsymbol{\Delta} = \begin{pmatrix} -1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Note that $\boldsymbol{\Delta}\mathbf{Y}_i$ contains the differences of the response from one visit to the next, so $\boldsymbol{\Delta}\mathbf{Y}_i$ is the response variable in model (2.10). The covariance matrix of the response for model (2.10) will then be $\boldsymbol{\Delta\Sigma\Delta}'$. Let us call $\mathbf{Z}$ the $r \times 2$ matrix of covariates for model (2.10),

$$\mathbf{Z}' = \begin{pmatrix} 1 & \cdots & 1 \\ k_i & \cdots & k_i \end{pmatrix}$$

and $\mathbf{X}_{[3,6]}$ a $(r+1) \times 2$ matrix containing the third and sixth column of $\mathbf{X}_i$ from model (2.9),

$$\mathbf{X}'_{[3,6]} = \begin{pmatrix} 0 & sj & sr \\ 0 & k_i sj & k_i sr \end{pmatrix}.$$

Then, it can be noted that $\frac{1}{s}\boldsymbol{\Delta}\mathbf{X}_{[3,6]} = \mathbf{Z}$. Therefore, the GLS estimate of λ_1 can be written as

$$\hat{\lambda}_1 = \left[\left(\sum_{i=1}^{N} \mathbf{Z}'_i \left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1} \mathbf{Z}_i\right)^{-1}\right]_{[2]} \left(\sum_{i=1}^{N} \mathbf{Z}'_i \left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1} \boldsymbol{\Delta}\mathbf{Y}_i\right) =$$

$$\left(\frac{1}{s}\sum_{i=1}^{N}\left[\left(\frac{1}{s^2}\sum_{i=1}^{N}\left(\boldsymbol{\Delta}\mathbf{X}_{[3,6]}\right)'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{X}_{[3,6]}\right)^{-1}\right]_{[2]}\left(\boldsymbol{\Delta}\mathbf{X}_{[3,6]}\right)'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{Y}_i\right)$$

$$= \left(\sum_{i=1}^{N} s\left[\left(\sum_{i=1}^{N}\mathbf{X}'_{[3,6]}\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{X}_{[3,6]}\right)^{-1}\right]_{[2]}\mathbf{X}'_{[3,6]}\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{Y}_i\right)$$

$$= \sum_{i=1}^{N}\mathbf{c}_\lambda \mathbf{Y}_i.$$

Now,

$$\sum_{i=1}^{N}\mathbf{X}'_{[3,6]}\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{X}_{[3,6]}$$

$$= s^2 \begin{pmatrix} 1 & \cdots & 1 \\ k_i & \cdots & k_i \end{pmatrix}_{(2\times r)} \left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}_{(r\times r)} \begin{pmatrix} 1 & k_i \\ \vdots & \vdots \\ 1 & k_i \end{pmatrix}_{(r\times 2)}$$

$$= s^2\left(\mathbf{t}'\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{t}\right)\begin{pmatrix} 1 & p_e \\ p_e & p_e \end{pmatrix}$$

so

$$\left[\left(\sum_{i=1}^{N}\mathbf{X}'_{[3,6]}\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{X}_{[3,6]}\right)^{-1}\right]_{[2]} =$$

$$\frac{1}{p_e(1-p_e)s^2\left(\mathbf{t}'\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{t}\right)}\begin{pmatrix} -p_e & 1 \end{pmatrix}$$

Now, by property B.3.5 of Seber (1984, page 536),

$$\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}.$$

Then,

$$\frac{1}{\left(\mathbf{t}'\boldsymbol{\Delta}'\left(\boldsymbol{\Delta\Sigma\Delta}'\right)^{-1}\boldsymbol{\Delta}\mathbf{t}\right)} = \frac{1}{\left(\mathbf{t}'\left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\right)\mathbf{t}\right)}$$

and with some algebra this expression equals

$$\frac{\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)}{\det(\mathbf{A})}\left(\right.$$

So

$$\left[\left(\sum_{i=1}^{N}\mathbf{X}'_{[3,6]}\mathbf{\Delta}'\left(\mathbf{\Delta\Sigma\Delta}'\right)^{-1}\mathbf{\Delta X}_{[3,6]}\right)^{-1}\right]_{[2]} = \frac{\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)}{p_e(1-p_e)s^2\det(\mathbf{A})}\begin{pmatrix} -p_e & 1 \end{pmatrix}$$

Now we need to derive $\mathbf{X}'_{[3,6]}\mathbf{\Delta}'\left(\mathbf{\Delta\Sigma\Delta}'\right)\mathbf{\Delta}$, and by using Seber's property again we have

$$\mathbf{X}'_{[3,6]}\mathbf{\Delta}'\left(\mathbf{\Delta\Sigma\Delta}'\right)\mathbf{\Delta} = \mathbf{X}'_{[3,6]}\left(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\mathbf{\Sigma}^{-1}\right).$$

So,

$$\begin{aligned}
\mathbf{c}_\lambda = s\left[\left(\sum_{i=1}^{N}\mathbf{X}'_{[3,6]}\mathbf{\Delta}'\left(\mathbf{\Delta\Sigma\Delta}'\right)^{-1}\mathbf{\Delta X}_{[3,6]}\right)^{-1}\right]_{[2]}\mathbf{X}'_{[3,6]}\mathbf{\Delta}'\left(\mathbf{\Delta\Sigma\Delta}'\right)\mathbf{\Delta} = \\
= \frac{\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)}{\det(\mathbf{A})sp_e(1-p_e)}\begin{pmatrix} -p_e & 1 \end{pmatrix}\mathbf{X}'_{[3,6]}\left(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\mathbf{\Sigma}^{-1}\right) \\
= \frac{\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)}{\det(\mathbf{A})p_e(1-p_e)}\begin{pmatrix} -p_e & 1 \end{pmatrix}\begin{pmatrix} \mathbf{t}' \\ k_i\mathbf{t}' \end{pmatrix}\left(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\mathbf{\Sigma}^{-1}\right) \\
= \frac{\left(-p_e + k_i\right)\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)}{\det(\mathbf{A})p_e(1-p_e)}\mathbf{t}'\left(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{1}\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}'\mathbf{\Sigma}^{-1}\right)
\end{aligned}$$

and we can observe that $\mathbf{c}_\lambda = \frac{\mathbf{c}_\eta}{s}$ and therefore $s\hat{\lambda}_1 = \hat{\eta}_5$ and $s^2Var\left(\hat{\lambda}_1\right) = Var\left(\hat{\eta}_5\right)$.

A.2 Bias and/or inefficiency of the ANCOVA and SLAIN tests in observational studies

Frison and Pocock (1992, 1997) considered general tests of the form

$$T = \frac{Np_e(1-p_e)\left(\bar{S}_1 - \bar{S}_0\right)^2}{\mathbf{c}'\mathbf{\Sigma c}},$$

where $\bar{S}_k$ is exposure group k 's mean, $k = (0,1)$, of a summary measure, S_i, that is a linear combination of the repeated measures of each subject, $S_i = \mathbf{c}'\mathbf{Y}_i$. The vector $\mathbf{c}'$ defines the summary measures, which could be, for example, the within-subject mean of the repeated measures, the within-subject slope, or ANCOVA, SLANC and SLAIN . Let n_k be the number of participants in exposure group k. Then,

$$\bar{S}_k = \frac{\sum_{i=1}^{N} S_i I\{k_i = k\}}{n_k} = \mathbf{c}' \left(\frac{\sum_{i=1}^{N} \mathbf{Y}_i I\{k_i = k\}}{n_k} \right) = \mathbf{c}'\bar{\mathbf{Y}}_k,$$

where $I\{k_i = k\}$ is an indicator variable that takes the value one when $k_i = k$ and zero otherwise, and $\bar{\mathbf{Y}}_k$ is the $(r+1) \times 1$ vector of sample means for each time in group k. Thus,

$$T = \frac{N p_e (1 - p_e) \left(\mathbf{c}' \left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0 \right) \right)^2}{\mathbf{c}' \boldsymbol{\Sigma} \mathbf{c}}.$$

Clearly, $\mathbb{E}\left[\bar{S}_k\right] = \mathbf{c}'\mu_k$, where μ_k is the vector of true means for each time in group k, and $\mathbb{E}\left[\bar{S}_1 - \bar{S}_0\right] = \mathbf{c}'(\mu_1 - \mu_0)$. If $\mathbf{c}' = \left(\frac{1}{r+1}, \cdots, \frac{1}{r+1}\right)$, we are testing the equality of the means of the two groups. Frison and Pocock (1992, 1997) found the vector $\mathbf{c}'$ so that $\mathbb{E}(T|H_0) = 0$ (valid) and for which the power of T is at its maximum possible under H_A (efficient). They found that the optimal vector $\mathbf{c}'$ is proportional to $(\mu_1 - \mu_0)\boldsymbol{\Sigma}^{-1}$. The resulting optimal test to detect a group difference under the CMD hypothesis in clinical trials was called "ANCOVA", and under CS covariance it has $\mathbf{c}' = \left(-\rho, \frac{1}{r}, \cdots, \frac{1}{r}\right)$ (Frison and Pocock, 1992, 1997). The resulting optimal test under the LDD hypothesis in clinical trials was called "SLAIN", and under CS covariance it has

$$c_j = \frac{12j + 6\rho r(2j - r - 1)}{r(r+1)\left[\rho r(r-1) + 2(2r+1)\right]},\ j = 0, \ldots, r$$

(Frison and Pocock, 1997). They also noted that their proof is similar to a GLS result (Frison and Pocock, 1997). Actually, since the GLS estimator is the best linear unbiased estimate of the parameter of a model, the test based on the GLS estimator of a model that correctly characterizes the shape of the differences between the exposed and the unexposed over time will be the optimal test. We next derive what is the underlying model of "ANCOVA" and "SLAIN".

Under the CMD hypothesis in clinical trials, the difference vector $(\mu_1 - \mu_0)$ is assumed to be proportional to $(0, 1, \ldots, 1)$ (Frison and Pocock, 1992), i.e. there is no difference at baseline, due to randomization, but there is a constant difference afterwards. This situation can be characterized by the following model:

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \beta_0 + \beta_1 I\left\{j > 0\right\} + \beta_2 I\left\{j > 0\right\} k_i, \tag{A.2}$$

where β_2 is the post-baseline difference between the two groups. The GLS estimator of the coefficients is

$$\hat{\mathbf{B}} = \left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{Y}_i\right) = \sum_{i=1}^{N} \left(\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i\right)^{-1} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{Y}_i\right).$$

The estimator of the parameter of interest is

$$\hat{\beta}_2 = \sum_{i=1}^{N} \left(\left[\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i\right)^{-1} \mathbf{X}'_i \mathbf{\Sigma}^{-1}\right]_{[3]} \mathbf{Y}_i\right),$$

where the subscript [3] refers to the third row of the matrix. We have

$$\begin{aligned}\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i &= \sum_{i=1}^{N} \left(\begin{array}{cccc} 1 & \cdots & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ 0 & k_i & \cdots & k_i \end{array}\right) \left(\begin{array}{ccc} v_{00} & & \\ \vdots & \ddots & \\ v_{0r} & \cdots & v_{rr} \end{array}\right) \left(\begin{array}{ccc} 1 & 0 & 0 \\ \vdots & 1 & k_i \\ \vdots & \vdots & \vdots \\ 1 & 1 & k_i \end{array}\right) \\ &= N \left(\begin{array}{ccc} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & & \\ \sum_{j=0}^{r}\sum_{j'=1}^{r} v_{jj'} & \sum_{j=1}^{r}\sum_{j'=1}^{r} v_{jj'} & \\ p_e \sum_{j=0}^{r}\sum_{j'=1}^{r} v_{jj'} & p_e \sum_{j=1}^{r}\sum_{j'=1}^{r} v_{jj'} & p_e \sum_{j=1}^{r}\sum_{j'=1}^{r} v_{jj'} \end{array}\right),\end{aligned}$$

where $v_{jj'}$ is the (j, j')th element of $\mathbf{\Sigma}^{-1}$. Let us call $a_1 = \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}$, $a_2 = \sum_{j=0}^{r}\sum_{j'=1}^{r} v_{jj'}$ and $a_3 = \sum_{j=1}^{r}\sum_{j'=1}^{r} v_{jj'}$. Then, the third row of $\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i\right)^{-1}$ is

$$\left(0, \;\; \frac{-1}{(1-p_e)a_3}, \;\; \frac{1}{p_e(1-p_e)a_3}\right).$$

We also have that

$$\mathbf{X'}_i\mathbf{\Sigma}^{-1} = \begin{pmatrix} \sum_{j=0}^{r} v_{j0} & \sum_{j=0}^{r} v_{j1} & \cdots & \sum_{j=0}^{r} v_{jr} \\ \sum_{j=1}^{r} v_{j0} & \sum_{j=1}^{r} v_{j1} & \cdots & \sum_{j=1}^{r} v_{jr} \\ k_i \sum_{j=1}^{r} v_{j0} & k_i \sum_{j=1}^{r} v_{j1} & \cdots & k_i \sum_{j=1}^{r} v_{jr} \end{pmatrix}$$

So, we can deduce

$$\left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \mathbf{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \mathbf{X'}_i \mathbf{\Sigma}^{-1} \right]_{[3]} =$$

$$\frac{1}{N} \left(\tfrac{1}{(1-p_e)a_3} \sum_{j=1}^{r} v_{j0} \left(-1 + \tfrac{k_i}{p_e}\right), \ \ldots, \ \tfrac{1}{(1-p_e)a_3} \sum_{j=1}^{r} v_{jr} \left(-1 + \tfrac{k_i}{p_e}\right) \right)$$

Then,

$$\hat{\beta}_2 = \sum_{i=1}^{N} \left(\left[\left(\sum_{i=1}^{N} \mathbf{X'}_i \mathbf{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \mathbf{X'}_i \mathbf{\Sigma}^{-1} \right]_{[3]} \mathbf{Y}_i \right) =$$

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} & \left(\left(\tfrac{1}{(1-p_e)a_3} \sum_{j=1}^{r} v_{j0} \left(-1 + \tfrac{k_i}{p_e}\right), \ \ldots, \ \tfrac{1}{(1-p_e)a_3} \sum_{j=1}^{r} v_{jr} \left(-1 + \tfrac{k_i}{p_e}\right) \right) \mathbf{Y}_i \right) \\ = & \left(\tfrac{-1}{a_3} \sum_{j=1}^{r} v_{j0}, \ \ldots, \ \tfrac{-1}{a_3} \sum_{j=1}^{r} v_{jr} \right) \frac{1}{N(1-p_e)} \sum_{i=1}^{N} \left(I\{k_i = 0\} \mathbf{Y}_i \right) \\ + & \left(\tfrac{1}{a_3} \sum_{j=1}^{r} v_{j0}, \ \ldots, \ \tfrac{1}{a_3} \sum_{j=1}^{r} v_{jr} \right) \frac{1}{Np_e} \sum_{i=1}^{N} \left(I\{k_i = 0\} \mathbf{Y}_i \right) \\ & = \left(\tfrac{1}{a_3} \sum_{j=1}^{r} v_{j0}, \ \ldots, \ \tfrac{1}{a_3} \sum_{j=1}^{r} v_{jr} \right) \left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0 \right). \end{aligned}$$

So, the $\mathbf{c}'$ vector is

$$\left(\tfrac{1}{a_3} \sum_{j=1}^{r} v_{j0}, \ \ldots, \ \tfrac{1}{a_3} \sum_{j=1}^{r} v_{jr} \right)$$

The inverse of a CS matrix has diagonal elements

$$\frac{1}{\sigma^2} \frac{1 + \rho(r-2) - \rho^2(r-1)}{(1-\rho)^2 (1 + r\rho)}$$

and off-diagonal elements

$$\frac{1}{\sigma^2}\frac{-\rho}{(1-\rho)\,(1+r\rho)}$$

(Graybill, 1983, theorem 8.3.4). Then, under CS,

$$a_3 = \sum_{j=1}^{r}\sum_{j'=1}^{r} v_{jj'} = \frac{r}{\sigma^2(1-\rho)(1+r\rho)},$$

$$\sum_{j=1}^{r} v_{j0} = \frac{-r\rho}{\sigma^2(1-\rho)(1+r\rho)}$$

and

$$\sum_{j=1}^{r} v_{jj'} = \frac{1}{\sigma^2(1-\rho)(1+r\rho)},\; j = 1,\ldots,r.$$

Therefore, $\hat{\beta}_2 = \left(-\rho,\ \frac{1}{r},\ \ldots,\ \frac{1}{r}\right)\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)$ and we can see that we get the same vector $\mathbf{c}' = \left(-\rho, \frac{1}{r}, \cdots, \frac{1}{r}\right)$ that Frison and Pocock (1992) derived for their "ANCOVA" analysis under CS.

Under the LDD hypothesis in clinical trials, the difference vector $(\mu_1 - \mu_0)$ is assumed to be proportional to $(0, 1, 2, \ldots, r)$, i.e. there is no difference at baseline, due to randomization, and afterwards the difference between the two groups changes linearly with time. This situation can be characterized by the following model:

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \beta_1 t_j + \beta_2 k_i t_j, \tag{A.3}$$

where β_2 is the difference in the rates of change in the two groups. The estimator of the parameter of interest is

$$\hat{\beta}_2 = \sum_{i=1}^{N}\left(\left[\left(\sum_{i=1}^{N}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)^{-1}\mathbf{X}'_i\boldsymbol{\Sigma}^{-1}\right]_{[2]}\mathbf{Y}_i\right),$$

where the subscript [2] refers to the second row of the matrix. We have

$$\sum_{i=1}^{N} \mathbf{X}'_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i = \sum_{i=1}^{N} \begin{pmatrix} 0 & j & r \\ 0 & k_i j & k_i r \end{pmatrix} \begin{pmatrix} v_{00} & & \\ \vdots & \ddots & \\ v_{0r} & \cdots & v_{rr} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ j & k_i j \\ r & k_i r \end{pmatrix}$$

$$= N \begin{pmatrix} \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'} & \\ p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'} & p_e \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'} \end{pmatrix},$$

where $v_{jj'}$ is the (j, j')th element of $\boldsymbol{\Sigma}^{-1}$. Then,

$$\left(\sum_{i=1}^{N} \mathbf{X}'_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} = \frac{1}{N p_e (1 - p_e) \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}} \begin{pmatrix} p_e & -p_e \\ -p_e & 1 \end{pmatrix}.$$

We can also derive $\mathbf{X}'_i \boldsymbol{\Sigma}^{-1}$. The j'th term ($j' = 0, \ldots, r$) in the first row of $\mathbf{X}'_i \boldsymbol{\Sigma}^{-1}$ has the form $\sum_{j=0}^{r} j v_{jj'}$, and the j'th term ($j' = 0, \ldots, r$) in the second row of $\mathbf{X}'_i \boldsymbol{\Sigma}^{-1}$ has the form $k_i \sum_{j=0}^{r} j v_{jj'}$. Thus, the j'th term ($j' = 0, \ldots, r$) in

$$\left[\left(\sum_{i=1}^{N} \mathbf{X}'_i \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \mathbf{X}'_i \boldsymbol{\Sigma}^{-1} \right]_{[2]}$$

has the expression

$$\frac{1}{N p_e (1 - p_e) \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}} \left(-p_e \sum_{j=0}^{r} j v_{jj'} + k_i \sum_{j=0}^{r} j v_{jj'} \right)$$

Then,

$$\hat{\beta}_2 = \sum_{i=1}^{N} \left(\left[\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \mathbf{X}'_i \mathbf{\Sigma}^{-1} \right]_{[2]} \mathbf{Y}_i \right) =$$

$$\frac{1}{N p_e (1-p_e) \sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}} \sum_{i=1}^{N} \sum_{j'=0}^{r} \left(-p_e \sum_{j=0}^{r} j v_{jj'} + k_i \sum_{j=0}^{r} j v_{jj'} \right) Y_{ij'}$$

$$= \frac{1}{\sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}}$$

$$\left\{ \sum_{j'=0}^{r} \left(- \sum_{j=0}^{r} j v_{jj'} \right) \frac{\sum_{i=1}^{N} \left(I\{k_i = 0\} Y_{ij'} \right)}{N(1-p_e)} + \sum_{j'=0}^{r} \left(\sum_{j=0}^{r} j v_{jj'} \right) \frac{\sum_{i=1}^{N} \left(I\{k_i = 1\} Y_{ij'} \right)}{N p_e} \right\}$$

$$= \frac{1}{\sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}} \left\{ \sum_{j'=0}^{r} \left(\sum_{j=0}^{r} j v_{jj'} \right) \bar{Y}_{1j'} - \sum_{j'=0}^{r} \left(\sum_{j=0}^{r} j v_{jj'} \right) \bar{Y}_{0j'} \right\}.$$

So, the j'th component of the $\mathbf{c}'$ vector is

$$\frac{\sum_{j=0}^{r} j v_{jj'}}{\sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'}}.$$

Under CS,

$$\sum_{j=0}^{r} j v_{jj'} = \frac{1}{\sigma^2 (1-\rho)(1+r\rho)} \left[j' \left(\frac{1 + \rho(r-2) - \rho^2 (r-1)}{(1-\rho)} \right) - \rho \sum_{j=0}^{r} j + \rho j' \right]$$
$$= \frac{1}{\sigma^2 (1-\rho)(1+r\rho)} \left[j' (1 + r\rho) - \frac{r(r+1)\rho}{2} \right]$$

and

$$\sum_{j=0}^{r} \sum_{j'=0}^{r} j j' v_{jj'} = \frac{r(r+1)(2 + r(4 + (r-1)\rho))}{12 \sigma^2 (1-\rho)(1+r\rho)}$$

Therefore, j'th component of the $\mathbf{c}'$ vector is

$$\frac{12 \left[j' (1 + r\rho) - \frac{r(r+1)\rho}{2} \right]}{r(r+1)(2 + r(4 + (r-1)\rho))} = \frac{12 j' + 6 \rho r (2j' - r - 1)}{r(r+1)(\rho r (r-1) + 2(2r+1))},$$

which is the expression that Frison and Pocock (1997) derived for CS.

In this paper we deal with observational studies, where the exposed and the unexposed already have a different expected value of the response at baseline. The CMD hypothesis in observational studies specifies that the mean group differences are constant over time, i.e. they are proportional to $(1, \ldots, 1)$. The mean difference between the two groups is constant at all time points and equal to $p_1\mu_{00}$, i.e. $(\mu_1 - \mu_0) = p_1\mu_{00}$. The parameter of interest under CMD is p_1, and one wants to test $H_0 : p_1 = 0$ vs. $H_A : p_1 \neq 0$. Under the CMD hypothesis for observational studies, the "ANCOVA" model is still unbiased under the null. We have $\mathbb{E}\left[\mathbf{c}'(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0)\middle| H_0\right] = \mathbf{c}'\mathbb{E}\left[(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0)\middle| H_0\right] = \mathbf{c}'\begin{pmatrix} p_1\mu_{00}, & \cdots & , p_1\mu_{00} \end{pmatrix}$. Since $p_1 = 0$ under H_0, we have $\mathbb{E}\left[\mathbf{c}'(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0)\middle| H_0\right] = 0$ for any $\mathbf{c}'$, and any vector $\mathbf{c}'$ produces unbiased estimators under the null. The optimal test is proportional to $(\mu_1 - \mu_0)\mathbf{\Sigma}^{-1}$ (Frison and Pocock, 1997), so in an observational study it is proportional to $(1, \ldots, 1)\mathbf{\Sigma}^{-1}$. Since ANCOVA is proportional to $(0, 1, \ldots, 1)\,\mathbf{\Sigma}^{-1}$, it is not the optimal test for CMD in observational studies.

The LDD hypothesis in observational studies specifies that the group mean differences are a linear function of time, but there is already a difference in the group means at baseline. In that case,

$$\mu_1 - \mu_0 = (p_1 + \frac{p_2 p_3}{\tau} t_j)\mu_{00}.$$

The parameter of interest under LDD is p_3, and one wants to test $H_0 : p_3 = 0$ vs. $H_A : p_3 \neq 0$. We have,

$$\begin{aligned} &\mathbb{E}\left[\mathbf{c}'(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0)\middle| H_0\right] = \mathbf{c}'\mathbb{E}\left[(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0)\middle| H_0\right] \\ &= \mathbf{c}'\begin{pmatrix} (p_1 + \frac{p_2 p_3}{\tau} t_0)\mu_{00}, & \cdots & , (p_1 + \frac{p_2 p_3}{\tau} t_r)\mu_{00} \end{pmatrix}'\Big|_{H_0 : p_3 = 0} = \mathbf{c}'\begin{pmatrix} p_1\mu_{00}, & \cdots & , p_1\mu_{00} \end{pmatrix}' \\ &\qquad = p_1\mu_{00}\mathbf{c}'\,(1, \ldots, 1)'. \end{aligned}$$

So, if there are differences at baseline, i.e. $p_1 \neq 0$, a test will be unbiased if and only if the sum of the components of $\mathbf{c}'$ is zero. For "SLAIN", the j'th component of $\mathbf{c}'$

is

$$\frac{\sum_{j=0}^{r} j v_{jj'}}{\sum_{j=0}^{r}\sum_{j'=0}^{r} j j' v_{jj'}},$$

so the test will be unbiased under the null if and only if $\sum_{j'=0}^{r}\sum_{j=0}^{r} j v_{jj'} = 0$. This will not be true in general. For example, under CS,

$$\sum_{j'=0}^{r}\sum_{j=0}^{r} j v_{jj'} = \frac{r(r+1)}{2\sigma^2 (1 + r\rho)}.$$

Therefore, "SLAIN" is biased in observational studies.

A.3 Proof that two-stage and GLS are equivalent approaches under CS or RS for $V(t_0) = 0$

In the setting where all subjects are observed at the same set of time points, this appendix will proof:

(i) That the estimator of the difference of the rates of change in the two exposure groups obtained using the summary measure (two-stage) approach is algebraically equivalent to the estimator of γ_3 obtained from fitting model (2.6) by OLS.

(ii) That when the covariance matrix $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$ has a CS or RS structure, the estimators from model (2.6) obtained by OLS and GLS are algebraically equivalent. Given (i), this implies that the estimator from the summary measure approach is algebraically equivalent to the GLS estimator. We also show that this is not the case for DEX.

Given (i) and (ii), since the estimators from the summary measure (two-stage) approach, and GLS are the same linear combination of $\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)$, $\mathbf{d}'\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)$, once we assume a covariance structure for $Var\left[\mathbf{Y}_i|\mathbf{X}_i\right] = \mathbf{\Sigma}_i$, the test statistic for the two methods is also equivalent and equal to

$$T = \frac{\mathbf{d}'\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)}{\sqrt{Var\left(\mathbf{d}'\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)\right)}},$$

where

$$\begin{aligned} Var\left(\mathbf{d}'\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)\right) &= \mathbf{d}' Var\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)\mathbf{d} \\ &= \mathbf{d}'\left(\frac{1}{Np_e} Var\left(\mathbf{Y}_{i,k_i=1}\right) + \frac{1}{N(1-p_e)} Var\left(\mathbf{Y}_{i,k_i=0}\right)\right)\mathbf{d} = \frac{\mathbf{d}'\mathbf{\Sigma}\mathbf{d}}{Np_e(1-p_e)}. \end{aligned}$$

Proof of (i)

Summary measure (two-stage) approach

Let $\mathbf{Z}_i$ be a $(r+1)\times 2$ matrix that contains a column of ones and the column of times for participant i. Since all subjects are observed at the same set of time points then $\mathbf{Z}_i = \mathbf{Z}$. Here, the summary measure is the subject-specific OLS slope associated with time from the regression of $\mathbf{Y}_i$ on $\mathbf{Z}_i = \mathbf{Z}$. Let us call $\hat{\beta}_i$, $i = 1,\dots,N$, the (2×1) vector containing the subject-specific intercept and slope of the regression, where $\hat{\beta}_i = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Y}_i$. The subject-specific intercepts and slopes are averaged in each exposure group as follows,

$$\hat{\beta}_k = \frac{\sum_{i=1}^{N}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Y}_i\, I\left\{k_i = k\right\}}{n_k} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\frac{\sum_{i=1}^{N}\mathbf{Y}_i I\left\{k_i = k\right\}}{n_k} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\bar{\mathbf{Y}}_k,$$

where $I\left\{k_i = k\right\}$ is an indicator variable that takes the value one when $k_i = k$ and zero otherwise; n_k is the number of participants in exposure group k, $k = 0, 1$; and $\bar{\mathbf{Y}}_k$ is the average of $\mathbf{Y}_i$ in group k. Since we are interested in the second component of $\hat{\beta}_k$, the slope associated with time, we define $\bar{S}_k = \left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\right)_{(2)}\bar{\mathbf{Y}}_k$, where the subscript (2) indicates the second row of the matrix $\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$. We are interested in the difference, which is $\left(\bar{S}_1 - \bar{S}_0\right) = \left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\right)_{(2)}\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)$.

OLS approach

With the OLS approach, we fit all the data at the same time, using

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \gamma_0 + \gamma_1 t_{ij} + \gamma_2 k_i + \gamma_3 \left(t_{ij} \times k_i\right),$$

and our interest in on γ_3. Reparameterizing, we can fit model

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \gamma_0^* \left(1 - k_i\right) + \gamma_1^* \left(1 - k_i\right) t_{ij} + \gamma_2^* k_i + \gamma_3^* k_i t_{ij},$$

and our parameter of interest is now $\gamma_3 = \gamma_3^* - \gamma_1^*$. The OLS estimator of the latter model can be derived as

$$\hat{\gamma}^* = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y} = \left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{Y}_i\right)$$

where $\mathbf{X}_i$ is the covariate matrix for subject i and can be written as $\mathbf{X}_i = \begin{pmatrix} \mathbf{Z} & \mathbf{0} \end{pmatrix}$ if participant i is unexposed and $\mathbf{X}_i = \begin{pmatrix} \mathbf{0} & \mathbf{Z} \end{pmatrix}$ if exposed. Then,

$$\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{X}_i\right) = \begin{pmatrix} N(1-p_e)\mathbf{Z}'\mathbf{Z} & \mathbf{0} \\ \mathbf{0} & Np_e\mathbf{Z}'\mathbf{Z} \end{pmatrix},$$

$$\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{X}_i\right)^{-1} = \frac{1}{N} \begin{pmatrix} \frac{1}{(1-p_e)}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{p_e}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \end{pmatrix}$$

and

$$\left(\sum_{i=1}^{N} \mathbf{X}'_i \mathbf{Y}_i\right) = \begin{pmatrix} \mathbf{Z}' \\ \mathbf{0} \end{pmatrix} \left(\sum_{i=1}^{N} \mathbf{Y}_i I\left\{k_i = 0\right\}\right) + \begin{pmatrix} \mathbf{0} \\ \mathbf{Z}' \end{pmatrix} \left(\sum_{i=1}^{N} \mathbf{Y}_i I\left\{k_i = 1\right\}\right)$$
$$= \begin{pmatrix} N(1-p_e)\mathbf{Z}'\bar{\mathbf{Y}}_0 \\ Np_e\mathbf{Z}'\bar{\mathbf{Y}}_1 \end{pmatrix}$$

so

$$\hat{\gamma}^* = \begin{pmatrix} \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\bar{\mathbf{Y}}_0 \\ \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\bar{\mathbf{Y}}_1 \end{pmatrix}$$

To compute $\hat{\gamma}_3 = \hat{\gamma}_3^* - \hat{\gamma}_1^*$ we need to subtract the second from the fourth component, so $\hat{\gamma}_3 = \left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\right)_{(2)}\left(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_0\right)$ as in the two-stage approach.

Proof of (ii)

A necessary and sufficient condition for the OLS and GLS estimators to be the same is $\mathbf{HV} = \mathbf{VH}$ (Puntanen and Styan, 1989, condition Z5), where $\mathbf{H}$ is the hat matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, $\mathbf{X}$ is our case the $N(r+1) \times 4$ matrix of covariates based on model (2.6), and $\mathbf{V}$ is the $N(r+1) \times N(r+1)$ covariance matrix of $\mathbf{Y}$, which is a block-diagonal matrix with the diagonal blocks equal to $\mathbf{\Sigma}$. As in the OLS derivation, we reparameterize the model as

$$\mathbb{E}\left(Y_{ij}|X_{ij}\right) = \gamma_0^* \left(1 - k_i\right) + \gamma_1^* \left(1 - k_i\right) t_{ij} + \gamma_2^* k_i + \gamma_3^* k_i t_{ij},$$

and for convenience we sort $\mathbf{X}$ sot that the first $N(1-p_e)$ participants are unexposed and therefore have $\mathbf{X}_i = \begin{pmatrix} \mathbf{Z} & \mathbf{0} \end{pmatrix}$, and the following Np_e are exposed and have $\mathbf{X}_i = \begin{pmatrix} \mathbf{0} & \mathbf{Z} \end{pmatrix}$. As derived in the OLS case,

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{N} \begin{pmatrix} \frac{1}{(1-p_e)} \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} & \mathbf{0}' \\ \mathbf{0} & \frac{1}{p_e} \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \end{pmatrix}$$

Then, it can be derived that

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{N} \begin{pmatrix} \mathbf{H}_{11} & \mathbf{0}' \\ \mathbf{0} & \mathbf{H}_{22} \end{pmatrix}$$

where $\mathbf{H}_{11}$ is a block matrix of $N(1-p_e) \times N(1-p_e)$ blocks, each block being equal to $\frac{1}{1-p_e}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$; and $\mathbf{H}_{22}$ is a block matrix with $Np_e \times Np_e$ blocks, each block being equal to $\frac{1}{p_e}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$. Since $\mathbf{V}$ is block diagonal with the diagonal blocks equal to $\mathbf{\Sigma}$, it follows that $\mathbf{HV}$ is going to be of the form

$$\mathbf{HV} = \frac{1}{N} \begin{pmatrix} (\mathbf{HV})_{11} & \mathbf{0}' \\ \mathbf{0} & (\mathbf{HV})_{22} \end{pmatrix}$$

where $(\mathbf{HV})_{11}$ is a block matrix of $N(1-p_e) \times N(1-p_e)$ blocks, each block being equal to $\frac{1}{1-p_e}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma}$; and $(\mathbf{HV})_{22}$ is a block matrix with $Np_e \times Np_e$ blocks, each block being equal to $\frac{1}{p_e}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma}$. Similarly, we can derive that $\mathbf{VH}$ is of the form

$$\mathbf{VH} = \frac{1}{N} \begin{pmatrix} (\mathbf{VH})_{11} & \mathbf{0}' \\ \mathbf{0} & (\mathbf{VH})_{22} \end{pmatrix}$$

where $(\mathbf{VH})_{11}$ is a block matrix of $N(1-p_e) \times N(1-p_e)$ blocks, each block being equal to $\frac{1}{1-p_e}\mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$; and $(\mathbf{VH})_{22}$ is a block matrix with $Np_e \times Np_e$ blocks, each block being equal to $\frac{1}{p_e}\mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$. Clearly, then, proving that $\mathbf{HV} = \mathbf{VH}$ is equivalent to proving that $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} = \mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$.

Next, we show that the $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} = \mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$ holds for $\mathbf{\Sigma}$ having a CS or RS structure and therefore the OLS and GLS estimators are algebraically equivalent in those cases. We also show that the condition does not hold for DEX.

CS

Under CS, $\mathbf{\Sigma} = \sigma^2\left(\rho\mathbf{1}\mathbf{1}' + (1-\rho)\mathbf{I}\right)$, where $\mathbf{I}$ is the $(r+1)\times(r+1)$ identity matrix and $\mathbf{1}$ a $(r+1)\times 1$ vector of ones. Then,

$$\begin{aligned}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} &= \sigma^2\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\left(\rho\mathbf{1}\mathbf{1}' + (1-\rho)\mathbf{I}\right)\\ &= \sigma^2\rho\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{1}\mathbf{1}' + \sigma^2(1-\rho)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'.\end{aligned}$$

Since $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$ is a projection matrix in the subspace defined by columns of $\mathbf{Z}$, and the first column of $\mathbf{Z}$is $\mathbf{1}$, then $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{1} = \mathbf{1}$ and $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} = \sigma^2\rho\mathbf{1}\mathbf{1}' + \sigma^2(1-\rho)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$. Now, we derive an expression for

$$\begin{aligned}\mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' &= \sigma^2\left(\rho\mathbf{1}\mathbf{1}' + (1-\rho)\mathbf{I}\right)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\\ &= \sigma^2\rho\mathbf{1}\mathbf{1}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' + \sigma^2(1-\rho)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'.\end{aligned}$$

For the same reasoning used above, $\mathbf{1}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \mathbf{1}'$, and therefore $\mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \sigma^2\rho\mathbf{1}\mathbf{1}' + \sigma^2(1-\rho)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$, which is the same expression we derived for $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma}$.

RS

Under RS, $\mathbf{\Sigma} = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{I}$. Then,

$$\begin{aligned}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} &= \mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\left(\mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{I}\right)\\ &= \mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'.\end{aligned}$$

Now, we derive an expression for

$$\mathbf{\Sigma}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \left(\mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{I}\right)\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}',$$

which is the same expression we derived for $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma}$.

DEX

A counterexample is enough to show that $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$ does not hold for DEX. With $r = 2$ then

$$\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix}.$$

If we take $\sigma^2 = 1$, $\rho = 0.8$ and $\theta = 1$ (AR(1) covariance structure) then

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 0.8 & 0.64 \\ 0.8 & 1 & 0.8 \\ 0.64 & 0.8 & 1 \end{pmatrix}.$$

Now,

$$\begin{aligned}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{\Sigma} &= \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix}\begin{pmatrix} 1 & 0.8 & 0.64 \\ 0.8 & 1 & 0.8 \\ 0.64 & 0.8 & 1 \end{pmatrix}\\ &= \begin{pmatrix} 0.993 & 0.866 & 0.633 \\ 0.813 & 0.866 & 0.813 \\ 0.633 & 0.866 & 0.993 \end{pmatrix}\end{aligned}$$

and

$$\mathbf{\Sigma Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}' = \begin{pmatrix} 1 & 0.8 & 0.64 \\ 0.8 & 1 & 0.8 \\ 0.64 & 0.8 & 1 \end{pmatrix} \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix} = \begin{pmatrix} 0.993 & 0.813 & 0.633 \\ 0.866 & 0.866 & 0.866 \\ 0.633 & 0.813 & 0.993 \end{pmatrix}.$$

We can see that the the [2,1], [1,2], [3,2] and [2,3] components differ, so the condition does not hold.

A.4 Proof that $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ is the same for $r = 1$ and $r = 2$ under LDD and $V(t_0) = 0$ with fixed follow-up period τ and equidistant time points.

When $V(t_0) = 0$, formula (3.4) is

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sum_{j=0}^{r}\left(\sum_{j'=0}^{r} v_{jj'}\right)r^2}{p_e(1-p_e)\tau^2 \det(\mathbf{A})},$$

where the term $v_{jj'}$ is the $[j, j']$ component of the inverse of $\mathbf{\Sigma}$ and

$$\mathbf{A} = \begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & \sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} \\ \sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & \sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ 0 & \cdots & r \end{pmatrix} \mathbf{\Sigma}^{-1} \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & r \end{pmatrix}.$$

Let

$$\mathbf{\Sigma}_1 = \begin{pmatrix} \sigma_{11} & \sigma_{1\tau} \\ \sigma_{1\tau} & \sigma_{\tau\tau} \end{pmatrix},$$

the covariance matrix when $r = 1$ and

$$\mathbf{\Sigma}_2 = \begin{pmatrix} \sigma_{11} & \sigma_{1,\tau/2} & \sigma_{1,\tau} \\ \sigma_{1,\tau/2} & \sigma_{\tau/2,\tau/2} & \sigma_{\tau/2,\tau} \\ \sigma_{1,\tau} & \sigma_{\tau/2,\tau} & \sigma_{\tau,\tau} \end{pmatrix}$$

when $r = 2$. Then, $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ will be the same for $r = 1$ and $r = 2$ if and only if

$$\frac{\mathbf{A}_1[1,1]}{\det(\mathbf{A}_1)} = \frac{4\mathbf{A}_2[1,1]}{\det(\mathbf{A}_2)},$$

where $\mathbf{A}_1$ is the $\mathbf{A}$ matrix when $r = 1$ and $\mathbf{A}_2$ is the $\mathbf{A}$ matrix when $r = 2$. We can now derive

$$\mathbf{\Sigma}_1^{-1} = \frac{1}{\sigma_{11}\sigma_{\tau\tau} - \sigma_{1\tau}^2} \begin{pmatrix} \sigma_{\tau\tau} & \\ -\sigma_{1\tau} & \sigma_{11} \end{pmatrix}$$

and

$$\mathbf{\Sigma}_2^{-1} = \frac{1}{-2\sigma_{1,\tau}\sigma_{1,\tau/2}\sigma_{\tau/2,\tau} + \sigma_{1,\tau}^2\sigma_{\tau/2,\tau/2} + \sigma_{1,\tau/2}^2\sigma_{\tau,\tau} + \sigma_{11}\left(\sigma_{\tau/2,\tau}^2 - \sigma_{\tau/2,\tau/2}\sigma_{\tau,\tau}\right)}$$
$$\begin{pmatrix} \sigma_{\tau/2,\tau}^2 - \sigma_{\tau/2,\tau/2}\sigma_{\tau,\tau} & & \\ \sigma_{1,\tau/2}\sigma_{\tau,\tau} - \sigma_{1,\tau}\sigma_{\tau/2,\tau} & \sigma_{1,\tau}^2 - \sigma_{11}\sigma_{\tau,\tau} & \\ \sigma_{1,\tau}\sigma_{\tau/2,\tau/2} - \sigma_{1,\tau/2}\sigma_{\tau/2,\tau} & \sigma_{11}\sigma_{\tau/2,\tau} - \sigma_{1,\tau}\sigma_{1,\tau/2} & \sigma_{1,\tau/2}^2 - \sigma_{11}\sigma_{\tau/2,\tau/2} \end{pmatrix}$$

Also,

$$\mathbf{A}_1 = \frac{1}{\sigma_{11}\sigma_{\tau\tau} - \sigma_{1\tau}^2} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sigma_{\tau\tau} & -\sigma_{1\tau} \\ -\sigma_{1\tau} & \sigma_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$
$$= \frac{1}{\sigma_{11}\sigma_{\tau\tau} - \sigma_{1\tau}^2} \begin{pmatrix} \sigma_{11} - 2\sigma_{1\tau} + \sigma_{\tau\tau} & \sigma_{11} - \sigma_{1\tau} \\ \sigma_{11} - \sigma_{1\tau} & \sigma_{11} \end{pmatrix}$$
$$\det(\mathbf{A}_1) = \frac{1}{\sigma_{11}\sigma_{\tau\tau} - \sigma_{1\tau}^2}$$

and

$$\frac{\mathbf{A}_1[1,1]}{\det(\mathbf{A}_1)} = \sigma_{11} - 2\sigma_{1\tau} + \sigma_{\tau\tau};$$

and

$$\mathbf{A}_2 = \frac{1}{-2\sigma_{1,\tau}\sigma_{1,\tau/2}\sigma_{\tau/2,\tau} + \sigma_{1,\tau}^2\sigma_{\tau/2,\tau/2} + \sigma_{1,\tau/2}^2\sigma_{\tau,\tau} + \sigma_{11}\left(\sigma_{\tau/2,\tau}^2 - \sigma_{\tau/2,\tau/2}\sigma_{\tau,\tau}\right)}$$
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$
$$\begin{pmatrix} \sigma_{\tau/2,\tau}^2 - \sigma_{\tau/2,\tau/2}\sigma_{\tau,\tau} & & \\ \sigma_{1,\tau/2}\sigma_{\tau,\tau} - \sigma_{1,\tau}\sigma_{\tau/2,\tau} & \sigma_{1,\tau}^2 - \sigma_{11}\sigma_{\tau,\tau} & \\ \sigma_{1,\tau}\sigma_{\tau/2,\tau/2} - \sigma_{1,\tau/2}\sigma_{\tau/2,\tau} & \sigma_{11}\sigma_{\tau/2,\tau} - \sigma_{1,\tau}\sigma_{1,\tau/2} & \sigma_{1,\tau/2}^2 - \sigma_{11}\sigma_{\tau/2,\tau/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$$

It can be derived that

$$\frac{\mathbf{A}_2[1,1]}{\det(\mathbf{A}_2)} = \frac{1}{\sigma_{11} + 2\sigma_{1,\tau} - 4\left(\sigma_{1,\tau/2} + \sigma_{\tau/2,\tau} - \sigma_{\tau/2,\tau/2}\right) + \sigma_{\tau,\tau}}$$
$$\left\{-\sigma_{1,\tau}^2 - \left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right)^2 + 2\sigma_{1,\tau}\left(\sigma_{1,\tau/2} + \sigma_{\tau/2,\tau} - \sigma_{\tau/2,\tau/2}\right) + \right.$$
$$\left.\left(\sigma_{\tau/2,\tau/2} - 2\sigma_{1,\tau/2}\right)\sigma_{\tau,\tau} + \sigma_{11}\left(\sigma_{\tau/2,\tau/2} + \sigma_{\tau,\tau} - 2\sigma_{\tau/2,\tau}\right)\right\}$$

Then,

$$\frac{\mathbf{A}_1[1,1]}{\det(\mathbf{A}_1)} = \frac{4\mathbf{A}_2[1,1]}{\det(\mathbf{A}_2)}$$

if and only if

$$\sigma_{11} - 2\sigma_{1\tau} + \sigma_{\tau\tau} = \frac{4}{\sigma_{11} + 2\sigma_{1,\tau} - 4\left(\sigma_{1,\tau/2} + \sigma_{\tau/2,\tau} - \sigma_{\tau/2,\tau/2}\right) + \sigma_{\tau,\tau}}$$
$$\left\{-\sigma_{1,\tau}^2 - \left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right)^2 + 2\sigma_{1,\tau}\left(\sigma_{1,\tau/2} + \sigma_{\tau/2,\tau} - \sigma_{\tau/2,\tau/2}\right) + \left(\sigma_{\tau/2,\tau/2} - 2\sigma_{1,\tau/2}\right)\sigma_{\tau,\tau} + \sigma_{11}\left(\sigma_{\tau/2,\tau/2} + \sigma_{\tau,\tau} - 2\sigma_{\tau/2,\tau}\right)\right\},$$

which with some algebra it reduces to $\sigma_{11} - \sigma_{\tau\tau} = 2\left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right)$. So, $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ will be the same for $r = 1$ and $r = 2$ if and only if $\sigma_{11} - \sigma_{\tau\tau} = 2\left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right)$. We can check that for the covariance structures used in the paper, i.e. compound symmetry (CS) (section 3.2), damped exponential (DEX) (section 3.3) and random intercepts and slopes (RS) (section 3.4) this condition is met. For CS,

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_{11} & \sigma_{1,\tau/2} & \sigma_{1,\tau} \\ \sigma_{1,\tau/2} & \sigma_{\tau/2,\tau/2} & \sigma_{\tau/2,\tau} \\ \sigma_{1,\tau} & \sigma_{\tau/2,\tau} & \sigma_{\tau,\tau} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

so $\sigma_{11} - \sigma_{\tau\tau} = \sigma^2(1-1) = 0$ and $2\left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right) = 2\sigma^2(\rho - \rho) = 0$ and the condition holds. For DEX,

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_{11} & \sigma_{1,\tau/2} & \sigma_{1,\tau} \\ \sigma_{1,\tau/2} & \sigma_{\tau/2,\tau/2} & \sigma_{\tau/2,\tau} \\ \sigma_{1,\tau} & \sigma_{\tau/2,\tau} & \sigma_{\tau,\tau} \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^{2^\theta} \\ \rho & 1 & \rho \\ \rho^{2^\theta} & \rho & 1 \end{pmatrix}$$

so $\sigma_{11} - \sigma_{\tau\tau} = \sigma^2(1-1) = 0$ and $2\left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right) = 2\sigma^2(\rho - \rho) = 0$ and the condition holds. For RS,

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma_{11} & \sigma_{1,\tau/2} & \sigma_{1,\tau} \\ \sigma_{1,\tau/2} & \sigma_{\tau/2,\tau/2} & \sigma_{\tau/2,\tau} \\ \sigma_{1,\tau} & \sigma_{\tau/2,\tau} & \sigma_{\tau,\tau} \end{pmatrix} =$$
$$\begin{pmatrix} \sigma_{b_0}^2 + \sigma_w^2 & & \\ \sigma_{b_0}^2 + \rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} & \sigma_{b_0}^2 + \sigma_{b_1}^2 + 2\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} + \sigma_w^2 & \\ \sigma_{b_0}^2 + 2\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} & \sigma_{b_0}^2 + 3\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} + 2\sigma_{b_1}^2 & \sigma_{b_0}^2 + 4\sigma_{b_1}^2 + 4\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} + \sigma_w^2 \end{pmatrix},$$

so

$$\sigma_{11} - \sigma_{\tau\tau} = \sigma_{b_0}^2 + \sigma_w^2 - \sigma_{b_0}^2 - 4\sigma_{b_1}^2 - 4\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} - \sigma_w^2 = -4\sigma_{b_1}^2 - 4\rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1}$$

and

$$2\left(\sigma_{1,\tau/2} - \sigma_{\tau/2,\tau}\right) = 2\left(\sigma_{b_0}^2 + \rho_{b_0 b_1}\sigma_{b_0}\sigma_{b_1} - \sigma_{b_0}^2 - 3\rho_{b_0 b_1}\sigma_{b_0}\sigma_{b_1} - 2\sigma_{b_1}^2\right) = -4\sigma_{b_1}^2 - 4\rho_{b_0 b_1}\sigma_{b_0}\sigma_{b_1}$$

and the condition holds.

A.5 Effect of p_e on r when $\mathbf{\Sigma}_i = \mathbf{\Sigma}$

When $\mathbf{\Sigma}_i = \mathbf{\Sigma}\ \forall i$, we can use equations (3.2) and (3.4) to write

$$N = \frac{(\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c})\left(z_\pi + z_{1-\alpha/2}\right)^2}{(\mathbf{c}'\mathrm{B}_{H_A})^2} = \frac{f(r)}{p_e(1-p_e)},$$

where $f(r)$ depends on r but not on p_e. We can define r implicitly as the value/s solving the equation $F(r) = 0$, where

$$F(r) = N - \frac{f(r)}{p_e(1-p_e)}.$$

Implicitly differentiating both sides of $F(r) = 0$, we obtain

$$\frac{\partial F(r)}{\partial p_e} = 0 \Leftrightarrow \frac{\partial r}{\partial p_e} = \frac{f(r)(1-2p_e)}{f'(r)p_e(1-p_e)}.$$

The value of p_e that minimizes r solves $\frac{\partial r}{\partial p_e} = 0$, and results in a single root, $p_e = 0.5$. Since $(1-2p_e) > 0$ for $p_e < 0.5$ and less than zero for $p_e > 0.5$, r has a maximum or a minimum at $p_e = 0.5$. The sign of $\frac{f(r)}{f'(r)}$ determines whether it is a maximum or a minimum. Since the variance $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ is always positive so is $f(r)$, and since the variance decreases as r increases, $f'(r)$ is negative. Therefore $\frac{f(r)}{f'(r)}$ is negative and $\frac{\partial r}{\partial p_e} < 0$ for $p_e < 0.5$ and $\frac{\partial r}{\partial p_e} > 0$ for $p_e > 0.5$, implying that r is minimum at $p_e = 0.5$.

A.6 When is there a limit to power less than 100% as $r \to \infty$?

A.6.1 CMD and CS

The inverse of a CS matrix has diagonal elements

$$\frac{1}{\sigma^2}\frac{1+\rho(r-2)-\rho^2(r-1)}{(1-\rho)^2\,(1+r\rho)}$$

and off-diagonal elements

$$\frac{1}{\sigma^2}\frac{-\rho}{(1-\rho)\,(1+r\rho)}$$

(Graybill, 1983, theorem 8.3.4). The sum of a row or a column of the inverse, $\sum_{j'=0}^{r} v_{jj'}$, is

$$\frac{1}{\sigma^2}\left(\frac{1+\rho(r-2)-\rho^2(r-1)}{(1-\rho)^2\,(1+r\rho)}-\frac{r\rho}{(1-\rho)\,(1+r\rho)}\right)=\frac{1}{\sigma^2\,(1+r\rho)}$$

and therefore

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}=\frac{r+1}{\sigma^2(1+r\rho)}.$$

Also,

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'}=\sum_{j=0}^{r} j\sum_{j'=0}^{r} v_{jj'}=\frac{r(r+1)}{2\sigma^2\,(1+r\rho)},$$

and

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'}=\frac{r(r+1)(2+r(4+(r-1)\rho))}{12\sigma^2(1-\rho)\,(1+r\rho)}.$$

Then,

$$\det(\mathbf{A})=\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}\right)\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'}\right)-\left(\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'}\right)^2=\frac{r(r+1)^2(r+2)}{12\sigma^4(1-\rho)\,(1+r\rho)}.$$

Plugging in these expressions in to equation (3.2), we have that under CMD and CS

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}=\frac{\sigma^2(1+r\rho)\left(r(r+2)(1+r\rho)s^2+12(1-\rho)V\left(t_0\right)\right)}{p_e(1-p_e)(r+1)\left(r(r+2)(1+r\rho)s^2+12(1-\rho)\left(1-\rho_{\mathrm{e},t_0}^2\right)V\left(t_0\right)\right)}.$$

Then, by comparing the highest order terms of r in the numerator and denominator of $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$, we can derive that

$$\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \lim_{r\to\infty} \frac{\sigma^2(1+r\rho)\left(r(r+2)(1+r\rho)s^2\right)}{p_e(1-p_e)(r+1)\left(r(r+2)(1+r\rho)s^2\right)} = \lim_{r\to\infty} \frac{\sigma^2(1+r\rho)}{p_e(1-p_e)(r+1)} = \frac{\sigma^2\rho}{p_e(1-p_e)}.$$

Since under CS, the covariance matrix of the response, $\mathbf{\Sigma}$, does not depend on s or τ, the results apply to both the fixed s and fixed τ design problems.

A.6.2 LDD and CS

Applying the results of Appendix A.6.1 to equation (3.4), we find that

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{12\sigma^2(1-\rho)(1-r\rho)}{p_e(1-p_e)(r+1)\left(r(r+2)(1+r\rho)\,s^2 + 12(1-\rho)\left(1-\rho^2_{\mathrm{e},t_0}\right)V(t_0)\right)}.$$

Since the denominator is a polynomial of fourth degree of r while the numerator is of first degree, then $\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = 0$.

Since under CS, the covariance matrix of the response, $\mathbf{\Sigma}$, does not depend on s or τ, the results apply to both the fixed s and fixed τ design problems.

A.6.3 CMD and AR(1)

A.6.3.1 When s is fixed

The AR(1) covariance matrix is given by (3.8) with $\theta = 1$, and its inverse is a tridiagonal matrix with the form

$$\mathbf{\Sigma}^{-1} = \frac{1}{(1-\rho^{2s})\,\sigma^2}\begin{pmatrix} 1 & -\rho^s & 0 & 0 & \cdots & 0 \\ -\rho^s & 1+\rho^{2s} & -\rho^s & 0 & & 0 \\ 0 & -\rho^s & 1+\rho^{2s} & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & -\rho^s & 0 \\ \vdots & & \ddots & -\rho^s & 1+\rho^{2s} & -\rho^s \\ 0 & 0 & \cdots & 0 & -\rho^s & 1 \end{pmatrix}$$

(Graybill, 1983, page 201). To use equation (3.2) we need expressions for $\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'}$, $\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'}$ and $\sum_{j=0}^{r}\sum_{j'=0}^{r} j j' v_{jj'}$. It can be easily shown that

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} = \frac{(1 + r + \rho^s - r\rho^s)}{\sigma^2 (1 + \rho^s)}.$$

Also,

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} = \frac{r(1-\rho^s)(1 + r(1-\rho^s) + \rho^s)}{2(1-\rho^{2s})\sigma^2}$$

and

$$\sum_{j=0}^{r}\sum_{j'=0}^{r} j j' v_{jj'} = \frac{r}{6(1-\rho^{2s})\sigma^2}\left(1 + 4\rho^s + \rho^{2s} + 3r\left(1-\rho^{2s}\right) + 2r^2(1-\rho^s)^2\right).$$

When $V(t_0) = 0$, we can use formula (3.3) to find

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sigma^2(1+\rho^s)}{p_e(1-p_e)(1 + r + \rho^s - r\rho^s)}$$

as given by Table 1. This formula has a first-order polynomial in r in the denominator, and has no terms involving r in the numerator, so $\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = 0$. If $V(t_0) > 0$, the formula is very long $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ and we used Mathematica (Wolfram Research Inc., 2005) to obtain the formula and compute the limit, which was zero. Therefore, $\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = 0$ when $V(t_0) > 0$.

A.6.3.2 When τ is fixed

When τ is fixed, we substitute s with τ/r in equation (3.2). So, when $V(t_0) = 0$,

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sigma^2(1+\rho^{\tau/r})}{p_e(1-p_e)(1 + r + \rho^{\tau/r} - r\rho^{\tau/r})},$$

and

$$\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \lim_{r\to\infty} \frac{\sigma^2(1+\rho^{\tau/r})}{p_e(1-p_e)\left[1 + \rho^{\tau/r} + r(1-\rho^{\tau/r})\right]}.$$

By l'Hôpital's rule, $\lim_{r\to\infty} r(1-\rho^{\tau/r}) = -\tau\log\rho$, and then

$$\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{2\sigma^2}{p_e(1-p_e)\left[2 - \tau\log\rho\right]}.$$

If $V(t_0) > 0$, we used Mathematica (Wolfram Research Inc., 2005) to derive the limit, which in this case has a very complicated expression,

$$2\sigma^2 \left(\left(\tau^3 + 12V(t_0)\tau\right)(\log(\rho))^2 - 6\left(\tau^2 + 4V(t_0)\right)\log(\rho) + 12\tau\right) [p_e(1-p_e)(2-\tau\log(\rho))]^{-1} \left\{\left(\tau^3 + 12V(t_0)\tau\right)(\log(\rho))^2 - 12V(t_0)(\tau\log(\rho)-2)\rho_{e,t_0}^2\log(\rho) - 6\left(\tau^2 + 4V(t_0)\right)\log(\rho) + 12\tau\right\}^{-1}$$

A.6.4 LDD and AR(1)

A.6.4.1 When s is fixed

Using the results from Appendix A.6.3 and applying formula (3.4) to the case $V(t_0) = 0$, we find that

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{12\sigma^2\left(1-\rho^{2s}\right)\left[r\, s^2 p_e(1-p_e)\right]^{-1}}{\left(2 + r(r+3) + 8\rho^s - 2r^2\rho^s + (r-2)(r-1)\rho^{2s}\right)}$$

as shown in Table 1. Since the denominator is a second degree polynomial of r while the numerator has no terms involving r, $\lim_{r\to\infty} \mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = 0$. If $V(t_0) > 0$, we used Mathematica (Wolfram Research Inc., 2005) and found that the limit was also zero.

A.6.4.2 When τ is fixed

When τ is fixed, we substitute s with τ/r. So, when $V(t_0) = 0$,

$$\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{12\sigma^2\left(1-\rho^{2\tau/r}\right) r\left[\tau^2 p_e(1-p_e)\right]^{-1}}{\left(2 + r(r+3) + 8\rho^{\tau/r} - 2r^2\rho^{\tau/r} + (r-2)(r-1)\rho^{2\tau/r}\right)},$$

as shown in Table 1. This expression can be rewritten as

$$\frac{12\sigma^2\left(1-\rho^{2\tau/r}\right)\left[p_e(1-p_e)\right]^{-1}}{\tau^2\left[r\left(1-\rho^{\tau/r}\right)^2 + \left(3\left(1-\rho^{2\tau/r}\right)\right) + \frac{1}{r}\left(2 + 8\rho^{\tau/r} + 2\rho^{2\tau/r}\right)\right]}.$$

Then, to compute $\lim_{r\to\infty} \mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ we note that the limit of the numerator is $12\sigma^2 \lim_{r\to\infty} (1-\rho^{2\tau/r}) = 0$. In the denominator, the limit of last two terms is zero, and l'Hôpital's rule can be used to show that the limit of the first term is also zero. We repeatedly applied l'Hôpital's rule to derive the limit of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$. With much algebra, we found that

$$\lim_{r\to\infty} \mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{24\sigma^2 \log\rho}{p_e(1-p_e)\left[-12\tau + 6\tau^2\log\rho - \tau^3(\log\rho)^2\right]}.$$

If $V(t_0) > 0$, using the expression derived for fixed s and substituting s by τ/r, we used Mathematica (Wolfram Research Inc., 2005) to derive the limit, which has a complicated expression,

$$24\sigma^2\log(\rho)\left[p_{\mathrm{e}}(1-p_{\mathrm{e}})\right] \left\{-\left(\tau^3 + 12V(t_0)\tau\right)\left(\log(\rho)\right)^2 + 12V(t_0)(\tau\log(\rho)-2)\rho^2_{e,t_0}\log(\rho) + 6\left(\tau^2 + 4V(t_0)\right)\log(\rho) - 12\tau\right\}^{-1}$$

A.6.5 CMD, RS and $V(t_0) = 0$

We only find the limit of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ when $V(t_0) = 0$. The covariance matrix of the repeated measurements is $\boldsymbol{\Sigma}_i = \mathbf{Z}_i\mathbf{D}\mathbf{Z}'_i + \sigma_w^2\mathbf{I}$, and since $V(t_0) = 0$, $\mathbf{Z}_i = \mathbf{Z}$ and $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma} = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{I}$. The matrix $\mathbf{Z}$ is $(r+1)\times 2$ and contains a column of ones and the column of times $(sj, j = 0,\ldots,r)$. Note that formula (3.2) depends on

$$\mathbf{A} = \begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & \sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} \\ \sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & \sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} \end{pmatrix}$$

only through $s^2\det(\mathbf{A})$. For convenience, we define a new matrix

$$\tilde{\mathbf{A}} = \begin{pmatrix} \sum_{j=0}^{r}\sum_{j'=0}^{r} v_{jj'} & s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} \\ s\sum_{j=0}^{r}\sum_{j'=0}^{r} j v_{jj'} & s^2\sum_{j=0}^{r}\sum_{j'=0}^{r} jj' v_{jj'} \end{pmatrix} = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z},$$

where $\det\left(\tilde{\mathbf{A}}\right) = s^2 \det(\mathbf{A})$. Then, $\tilde{\mathbf{A}} = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z} = \mathbf{Z}'\left(\mathbf{ZDZ}' + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{Z}$. Using the property

$$(\mathbf{GBG}' + \mathbf{C})^{-1} = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}\left(\mathbf{B}^{-1} + \mathbf{G}'\mathbf{C}^{-1}\mathbf{G}\right)^{-1}\mathbf{G}'\mathbf{C}^{-1},$$

which can be found in (Timm, 2002, property 8, page 46), we have that

$$\begin{aligned}\left(\mathbf{ZDZ}' + \sigma_w^2\mathbf{I}\right)^{-1} &= \frac{1}{\sigma_w^2}\mathbf{I} - \frac{1}{\sigma_w^2}\mathbf{IZ}\left(\mathbf{D}^{-1} + \mathbf{Z}'\frac{1}{\sigma_w^2}\mathbf{IZ}\right)^{-1}\mathbf{Z}'\mathbf{I}\frac{1}{\sigma_w^2} \\ &= \frac{1}{\sigma_w^2}\mathbf{I} - \frac{1}{\sigma_w^4}\mathbf{Z}\left(\mathbf{D}^{-1} + \frac{1}{\sigma_w^2}\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'.\end{aligned}$$

Now,

$$\begin{aligned}&\mathbf{Z}'\left(\mathbf{ZDZ}' + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{Z} \\ &\qquad = \frac{1}{\sigma_w^2}\mathbf{Z}'\mathbf{Z} - \frac{1}{\sigma_w^4}\mathbf{Z}'\mathbf{Z}\left(\mathbf{D}^{-1} + \frac{1}{\sigma_w^2}\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Z}\end{aligned}$$

and using the property

$$\mathbf{G}^{-1} - \mathbf{G}^{-1}\left(\mathbf{G}^{-1} + \mathbf{B}^{-1}\right)^{-1}\mathbf{G}^{-1} = (\mathbf{G} + \mathbf{B})^{-1},$$

(Timm, 2002, property 6, page 46),

$$\begin{aligned}\tilde{\mathbf{A}} &= \mathbf{Z}'\left(\mathbf{ZDZ}' + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{Z} \\ &= \frac{1}{\sigma_w^2}\mathbf{Z}'\mathbf{Z} - \frac{1}{\sigma_w^4}\mathbf{Z}'\mathbf{Z}\left(\mathbf{D}^{-1} + \frac{1}{\sigma_w^2}\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Z} = \left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\sigma_w^2 + \mathbf{D}\right)^{-1} \\ &= \left(\sigma_w^2\begin{pmatrix}\frac{2+4r}{(r+1)(r+2)} & \frac{-6}{s(r+1)(r+2)} \\ \frac{-6}{s(r+1)(r+2)} & \frac{12}{s^2r(r+1)(r+2)}\end{pmatrix} + \begin{pmatrix}\sigma_{b_0}^2 & \rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} \\ \rho_{b_0b_1}\sigma_{b_0}\sigma_{b_1} & \sigma_{b_1}^2\end{pmatrix}\right)^{-1}.\end{aligned}$$

We computed this inverse with Mathematica (Wolfram Research Inc., 2005), and then by substituting $s^2\det(\mathbf{A})$ by $\det\left(\tilde{\mathbf{A}}\right)$ into equation (3.3), we found $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$, which is

$$\frac{\left(\sigma_0^2 + \frac{2(2r+1)\sigma_w^2}{r^2+3r+2}\right)\left(\sigma_1^2 + \frac{12\sigma_w^2}{(r^3+3r^2+2r)s^2}\right) - \left(\sigma_{01} - \frac{6\sigma_w^2}{(r^2+3r+2)s}\right)^2}{(1-p_e)p_e\left(\sigma_1^2 + \frac{12\sigma_w^2}{(r^3+3r^2+2r)s^2}\right)}.$$

The limit of this expression is

$$\frac{\sigma_0^2\sigma_1^2 - \sigma_{01}^2}{p_e(1-p_e)\sigma_1^2}.$$

The same result is obtained when τ is fixed. This limit can be rewritten as

$$\lim_{r\to\infty} \ \mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \frac{\sigma_{t_0}^2\rho_{t_0}\left(1-\rho_{01}^2\right)}{p_e(1-p_e)}.$$

When $V(t_0) > 0$ and the covariance follows RS, the full distribution of (k,t) is needed and numerical integration needs to be performed (Appendix A.8). Thus, general results about the limit of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ as $r \to \infty$ cannot be obtained.

A.6.6 LDD, RS and $V(t_0) = 0$

When $V(t_0) = 0$, using equation (3.4), we can derive $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ by substituting $s^2\det(\mathbf{A})$ by $\det\left(\tilde{\mathbf{A}}\right)$ to obtain, in terms of our parameterization,

$$\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2p_e(1-p_e)}\right)\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right)$$

as in Table 1. Then,

$$\lim_{r\to\infty} \ \mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c} = \left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2p_e(1-p_e)}\right)\left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)},$$

and, equivalently when τ is fixed, the limit is

$$\left(\frac{12\sigma^2(1-\rho_{t_0})}{\tau^2p_e(1-p_e)}\right)\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}.$$

When $V(t_0) > 0$ and the covariance follows RS, the full distribution of (k,t) is needed and numerical integration needs to be performed (Appendix A.8). Thus, general results about the limit of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ as $r \to \infty$ cannot be obtained.

A.7 The effect of covariance parameters on the minimum r for a fixed N, subject to power π

A.7.1 The effect of ρ and ρ_{t_0}

A.7.1.1 CMD, CS, $V(t_0) = 0$

From equation (3.6),

$$r = \frac{\beta_2^2 N\, p_e(1-p_e) - \left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2}{\left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2 \rho - \beta_2^2 N\, p_e(1-p_e)}.$$

Differentiating with respect to ρ, we get

$$\frac{\partial r}{\partial \rho} = \frac{\left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2 \left(-\beta_2^2 N\, p_e(1-p_e) + \left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2\right)}{\left(\left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2 \rho - \beta_2^2 N\, p_e(1-p_e)\right)^2}.$$

If $\left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2 > \beta_2^2 N\, p_e(1-p_e)$, then $\frac{\partial r}{\partial \rho} > 0$, so r increases as ρ increases. If $\left(z_\pi + z_{1-\alpha/2}\right)^2 \sigma^2 < \beta_2^2 N\, p_e(1-p_e)$, then $\frac{\partial r}{\partial \rho} < 0$, so r decreases as ρ increases.

A.7.1.2 LDD, CS, fixed s, $V(t_0) = 0$

The minimum r for fixed N and fixed power, π, solves

$$N = \frac{12\sigma^2(1-\rho)\left(z_\pi + z_{1-\alpha/2}\right)^2}{\gamma_3^2 p_e(1-p_e)\; s^2\, r\,(r+1)(r+2)},$$

which was obtained plugging in the corresponding value of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F\left(r,\, \rho\right) = \frac{N\gamma_3^2 p_e(1-p_e)\; s^2}{12\sigma^2\left(z_\pi + z_{1-\alpha/2}\right)^2} - \frac{(1-\rho)}{r\,(r+1)(r+2)},$$

the equation $F\left(r,\, \rho\right) = 0$ implicitly defines the function $r = f(\rho)$. Using implicit differentiation and taking into account that r is a function of ρ, $r(\rho)$, we obtain

$$\frac{\partial r}{\partial \rho} = \frac{-r\ (r+1)\,(r+2)}{(1-\rho)(3r^2+6r+2)}.$$

Since r is positive, the derivative is always negative, and r decreases as ρ increases.

A.7.1.3 LDD, CS, fixed τ, $V(t_0) = 0$

The minimum r for fixed N and fixed power, π, solves

$$N = \frac{12\sigma^2(1-\rho)\left(z_\pi + z_{1-\alpha/2}\right)^2 r}{\gamma_3^2 p_e(1-p_e)\,\tau^2\,(r+1)(r+2)},$$

which was obtained plugging in the corresponding value of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F(r,\,\rho) = \frac{N\gamma_3^2 p_e(1-p_e)\,\tau^2}{12\sigma^2\left(z_\pi + z_{1-\alpha/2}\right)^2} - \frac{(1-\rho)r}{(r+1)(r+2)},$$

the equation $F(r,\,\rho) = 0$ implicitly defines the function $r = f(\rho)$. Using implicit differentiation and taking into account that r is a function of ρ, we obtain

$$\frac{\partial r}{\partial \rho} = \frac{r\,(r+1)\,(r+2)}{(1-\rho)\,(-r^2+2)}.$$

If $r \geqslant 2$, then $\frac{\partial r}{\partial \rho} < 0$. So if we are taking at least two post-baseline measures, larger values of ρ lead to smaller values of r to achieve the specified power. Since $\frac{r}{(r+1)(r+2)}$ is the same for $r = 1$ and $r = 2$, it is preferable to choose $r = 1$ since fewer measurements need to be collected. Therefore, the choice between $r = 1$ and $r = 2$ is not affected by ρ.

A.7.1.4 LDD, RS, fixed s, $V(t_0) = 0$

The minimum r for fixed N and fixed power, π, solves

$$N = \frac{\left(z_\pi + z_{1-\alpha/2}\right)^2 \left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2}\right)\left(\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right)\right.}{\gamma_3^2 p_e\left(1-p_e\right)},$$

which was obtained by plugging in the corresponding value of $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F(r,\,\rho_{t_0}) = \frac{N\gamma_3^2 p_e(1-p_e)s^2}{12\sigma^2\left(z_\pi + z_{1-\alpha/2}\right)^2} - (1-\rho_{t_0})\left(\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right),\right.$$

the equation $F\left(r,\ \rho_{t_0}\right)=0$ implicitly defines the function $r=f(\rho_{t_0})$. Using implicit differentiation and taking into account that r depends on ρ_{t_0}, we obtain

$$\frac{\partial r}{\partial \rho_{t_0}}=\frac{-r\ (r+1)\,(r+2)\left[\rho_{b_1,s,\tilde{r}} r\ (r+1)\,(r+2)+\left(1-\rho_{b_1,s,\tilde{r}}\right)\tilde{r}(\tilde{r}+1)(\tilde{r}+2)\right]}{\left(1-\rho_{b_1,s,\tilde{r}}\right)\tilde{r}(\tilde{r}+1)(\tilde{r}+2)(1-\rho_{t_0})\left(3r^2+6r+2\right)}<0.$$

Since the derivative is always negative when $r>0$, r decreases as ρ_{t_0} increases.

A.7.1.5 LDD, RS, fixed τ, $V(t_0)=0$

The minimum r for fixed N and fixed power, π, solves

$$N=\frac{\left(z_\pi+z_{1-\alpha/2}\right)^2\left(\frac{12\sigma^2(1-\rho_{t_0})}{\tau^2}\right)\left(\frac{r}{(r+1)(r+2)}+\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)}{\gamma_3^2 p_e(1-p_e)},$$

which was obtained by plugging in the corresponding value of $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F\left(r,\ \rho\right)=$$
$$\frac{N\gamma_3^2 p_e(1-p_e)\tau^2}{12\sigma^2\left(z_\pi+z_{1-\alpha/2}\right)^2}-(1-\rho_{t_0})\left(\frac{r}{(r+1)(r+2)}+\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)$$

the equation $F\left(r,\ \rho_{t_0}\right)=0$ implicitly defines the function $r=f(\rho_{t_0})$. Using implicit differentiation, and taking into account that r depends on ρ_{t_0}, we obtain

$$\frac{\partial r}{\partial \rho_{t_0}}=\frac{(r+1)\,(r+2)\left[\rho_{b_1,\tau,\tilde{r}}\tilde{r}\,(r+1)\,(r+2)+\left(1-\rho_{b_1,\tau,\tilde{r}}\right)(\tilde{r}+1)(\tilde{r}+2)r\right]}{\left(1-\rho_{b_1,\tau,\tilde{r}}\right)(\tilde{r}+1)(\tilde{r}+2)\,(1-\rho_{t_0})\left(2-r^2\right)}.$$

If $r\geqslant 2$ then $\frac{\partial r}{\partial \rho_{t_0}}<0$. So if we are taking at least two post-baseline measures, larger values of ρ_{t_0} lead to smaller minimal values of r to achieve a certain power. Since $\frac{r}{(r+1)(r+2)}$ is the same for $r=1$ and $r=2$, the resulting power of both studies would be the same and it would be preferable to choose $r=1$ since less measurements need to be collected. The choice between $r=1$ and $r=2$ is not affected by ρ_{t_0}.

A.7.2 The effect of $\rho_{b_1,s,\tilde{r}}$

A.7.2.1 LDD, RS, fixed s, $V(t_0) = 0$

The minimum r for fixed N and fixed power, π, solves

$$N = \frac{\left(z_\pi + z_{1-\alpha/2}\right)^2 \left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2}\right)\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right)}{\gamma_3^2 p_e(1-p_e)},$$

which was obtained plugging in the corresponding value of $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F\left(r,\ \rho_{b_1,s,\tilde{r}}\right) = \frac{N\gamma_3^2 p_e(1-p_e)s^2}{12\sigma^2\left(z_\pi + z_{1-\alpha/2}\right)^2(1-\rho_{t_0})} - \left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right),$$

the equation $F\left(r,\ \rho_{b_1,s,\tilde{r}}\right) = 0$ implicitly defines the function $r = f(\rho_{b_1,s,\tilde{r}})$. Using implicit differentiation, and taking into account that r depends on $\rho_{b_1,s,\tilde{r}}$, we obtain

$$\frac{\partial r}{\partial \rho_{b_1,s,\tilde{r}}} = \frac{r^2(r+1)^2(r+2)^2}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)(1-\rho_{b_1,s,\tilde{r}})^2(3r^2+6r+2)} > 0.$$

Since the derivative is always positive, r increases as $\rho_{b_1,s,\tilde{r}}$ increases.

A.7.2.2 LDD, RS, fixed τ, $V(t_0) = 0$

The minimum r for fixed N and fixed power, π, solves

$$N = \frac{\left(z_\pi + z_{1-\alpha/2}\right)^2 \left(\frac{12\sigma^2(1-\rho_{t_0})}{\tau^2}\right)\left(\frac{r}{(r+1)(r+2)} + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)}{\gamma_3^2 p_e(1-p_e)},$$

which was obtained plugging in the corresponding value of $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ in Table 1 into equation (3.5). Defining

$$F\left(r,\ \rho_{b_1,\tau,\tilde{r}}\right) = \frac{N\gamma_3^2 p_e(1-p_e)\tau^2}{12\sigma^2\left(z_\pi + z_{1-\alpha/2}\right)^2(1-\rho_{t_0})} - \left(\frac{r}{(r+1)(r+2)} + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)\Bigg($$

the equation $F\left(r,\, \rho_{b_1,s,\tilde{r}}\right) = 0$ implicitly defines the function $r = f(\rho_{b_1,s,\tilde{r}})$. Using implicit differentiation, and taking into account that r depends on $\rho_{b_1,\tau,\tilde{r}}$, we obtain

$$\frac{\partial}{\partial \rho_{b_1,\tau,\tilde{r}}} = \frac{\tilde{r}(r+1)^2(r+2)^2}{(r^2-2)(\tilde{r}+1)(\tilde{r}+2)(1-\rho_{b_1,\tau,\tilde{r}})^2}.$$

If $r \geqslant 2$, $\frac{\partial r}{\partial \rho_{b_1,\tau,\tilde{r}}} > 0$. So if we are taking at least two post-baseline measurements, the effect of increasing $\rho_{b_1,\tau,\tilde{r}}$ is to increase the minimum r needed to achieve a pre-specified power. Since $\frac{r}{(r+1)(r+2)}$ is the same for $r = 1$ and $r = 2$, the resulting power of both studies would be the same and it is therefore preferable to choose $r = 1$ since less measurements need to be collected. The choice between $r = 1$ and $r = 2$ is not affected by $\rho_{b_1,\tau,\tilde{r}}$.

A.8 Calculation of $\boldsymbol{\Sigma}_\mathrm{B}$ under RS and $V(t_0) > 0$

We need to derive

$$\boldsymbol{\Sigma}_\mathrm{B} = \left(\mathbb{E}\left(\mathbf{X'}_i \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i\right)\right)^{-1}.$$

When $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$ for all subjects, $\mathbb{E}\left(\mathbf{X'}_i \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i\right)$ can be computed exactly. Under RS, $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$ when $V(t_0) = 0$, in which case equations (3.3) and (3.4) with $V(t_0) = 0$ provide expressions for $\mathbf{c}'\boldsymbol{\Sigma}_\mathrm{B}\mathbf{c}$ for CMD and LDD, respectively. However, if $V(t_0) > 0$, then $\boldsymbol{\Sigma}_i \neq \boldsymbol{\Sigma}$ under RS. Specifically, $\boldsymbol{\Sigma}_i$ depends on t_{0i}, so we have $\boldsymbol{\Sigma}(t_{0i})$. The formula for $\boldsymbol{\Sigma}_i$ under RS is $\boldsymbol{\Sigma}_i = \mathbf{Z}_i \mathbf{D} \mathbf{Z'}_i + \sigma_w^2 \mathbf{I}$, where

$$\mathbf{Z'}_i = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 \\ t_{0i} & \cdots & t_{0i}+js & \cdots & t_{0i}+rs \end{pmatrix}$$

and

$$\mathbf{D} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$$

Under LDD,

$$\mathbf{Z}_i\mathbf{W}_i = \begin{pmatrix} 1 & t_{0i} \\ \vdots & \vdots \\ \vdots & t_{0i}+js \\ \vdots & \vdots \\ 1 & t_{0i}+rs \end{pmatrix} \begin{pmatrix} 1 & 0 & k_i & 0 \\ 0 & 1 & 0 & k_i \end{pmatrix} =$$

$$\begin{pmatrix} 1 & t_{0i} & k_i & t_{0i}k_i \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & t_{0i}+js & \vdots & (t_{0i}+js)\,k_i \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{0i}+rs & k_i & (t_{0i}+rs)\,k_i \end{pmatrix} = \mathbf{X}_i,$$

where

$$\mathbf{W}_i = \begin{pmatrix} 1 & 0 & k_i & 0 \\ 0 & 1 & 0 & k_i \end{pmatrix}.$$

Therefore,

$$\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i = \mathbf{W}_i'\mathbf{Z}_i'\mathbf{\Sigma}_i^{-1}\mathbf{Z}_i\mathbf{W}_i = \mathbf{W}_i'\mathbf{Z}_i'\left(\mathbf{Z}_i\mathbf{D}\mathbf{Z}'_i + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{Z}_i\mathbf{W}_i.$$

In Appendix A.6.5, we showed that

$$\mathbf{Z}'\left(\mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{Z} = \left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\sigma_w^2 + \mathbf{D}\right)^{-1},$$

so $\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i = \mathbf{W}_i'\left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\sigma_w^2 + \mathbf{D}\right)^{-1}\mathbf{W}_i$. Now,

$$(\mathbf{Z}'_i\mathbf{Z}_i) = \begin{pmatrix} r+1 & (r+1)\,t_{0i} + \frac{s\,r(r+1)}{2} \\ (r+1)\,t_{0i} + \frac{s\,r(r+1)}{2} & (r+1)\,t_{0i}^2 + st_{0i}r(r+1) + \frac{s^2r(r+1)(2r+1)}{6} \end{pmatrix}$$

$$(\mathbf{Z}'_i\mathbf{Z}_i)^{-1} = \frac{2}{r(r+1)(r+2)s^2}\begin{pmatrix} r\,(1+2r)\,s^2 + 6r\,st_{0i} + 6t_{0i}^2) & -3\,(r\,s+2t_{0i}) \\ -3\,(r\,s+2t_{0i}) & 6 \end{pmatrix}$$

and

$$\left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\sigma_w^2 + \mathbf{D}\right)^{-1} = \begin{pmatrix} a(t_{0i}) & c(t_{0i}) \\ c(t_{0i}) & d(t_{0i}) \end{pmatrix},$$

where

$$a(t_{0i}) = \frac{\left(\frac{12\sigma_w^2}{r(r+1)(r+2)s^2} + \sigma_1^2\right)}{\left(\frac{12\sigma_w^2}{r(r+1)(r+2)s^2} + \sigma_1^2\right)\left(\sigma_0^2 + \frac{2\sigma_w^2\left(r(1+2r)s^2+6rst_{0i}+6t_{0i}^2\right)}{r(r+1)(r+2)s^2}\right) - \left(\sigma_{01} - \frac{6\sigma_w^2(rs+2t_{0i})}{r(r+1)(r+2)s^2}\right)^2},$$

$$c(t_{0i}) = \frac{\left(-\sigma_{01} + \frac{6\sigma_w^2(rs+2t_{0i})}{r(r+1)(r+2)s^2}\right)}{\left(\frac{12\sigma_w^2}{r(r+1)(r+2)s^2} + \sigma_1^2\right)\left(\sigma_0^2 + \frac{2\sigma_w^2\left(r(1+2r)s^2+6rst_{0i}+6t_{0i}^2\right)}{r(r+1)(r+2)s^2}\right) - \left(\sigma_{01} - \frac{6\sigma_w^2(rs+2t_{0i}}{r(r+1)(r+2)s^2}\right)^2},$$

and

$$d(t_{0i}) = \frac{\left(\sigma_0^2 + \frac{2\sigma_w^2\left(r(1+2r)s^2+6rst_{0i}+6t_{0i}^2\right)}{r(r+1)(r+2)s^2}\right)}{\left(\frac{12\sigma_w^2}{r(r+1)(r+2)s^2} + \sigma_1^2\right)\left(\sigma_0^2 + \frac{2\sigma_w^2\left(r(1+2r)s^2+6rst_{0i}+6t_{0i}^2\right)}{r(r+1)(r+2)s^2}\right) - \left(\sigma_{01} - \frac{6\sigma_w^2(rs+2t_{0i}}{r(r+1)(r+2)s^2}\right)^2}.$$

Pre- and post-multiplying by $\mathbf{W}_i$, we obtain

$$\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i = \mathbf{W}_i'\left(\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\sigma_w^2 + \mathbf{D}\right)^{-1}\mathbf{W}_i = \begin{pmatrix} a(t_{0i}) & c(t_{0i}) & k_i a(t_{0i}) & k_i c(t_{0i}) \\ c(t_{0i}) & d(t_{0i}) & k_i c(t_{0i}) & k_i d(t_{0i}) \\ k_i a(t_{0i}) & k_i c(t_{0i}) & k_i a(t_{0i}) & k_i c(t_{0i}) \\ k_i c(t_{0i}) & k_i d(t_{0i}) & k_i c(t_{0i}) & k_i d(t_{0i}) \end{pmatrix},$$

since $k_i^2 = k_i$. To compute

$$\mathbb{E}\left[\mathbf{X}_i'\mathbf{\Sigma}_i^{-1}\mathbf{X}_i\right] = \begin{pmatrix} \mathbb{E}[a(t_{0i})] & \mathbb{E}[c(t_{0i})] & \mathbb{E}[k_i a(t_{0i})] & \mathbb{E}[k_i c(t_{0i})] \\ \mathbb{E}[c(t_{0i})] & \mathbb{E}[d(t_{0i})] & \mathbb{E}[k_i c(t_{0i})] & \mathbb{E}[k_i d(t_{0i})] \\ \mathbb{E}[k_i a(t_{0i})] & \mathbb{E}[k_i c(t_{0i})] & \mathbb{E}[k_i a(t_{0i})] & \mathbb{E}[k_i c(t_{0i})] \\ \mathbb{E}[k_i c(t_{0i})] & \mathbb{E}[k_i d(t_{0i})] & \mathbb{E}[k_i c(t_{0i})] & \mathbb{E}[k_i d(t_{0i})] \end{pmatrix},$$

the distribution of t_{0i} and the joint distribution of (t_{0i}, k_i) are needed. Assuming that k_i follows a Bernoulli distribution with probability of success p_e, we have

$$\begin{aligned}\mathbb{E}[a(t_{0i})] &= \int a(t_{0i}) f(t_{0i}) dt_{0i} \\ &= (1-p_e)\int a(t_{0i}) f(t_{0i}|k_i=0) dt_{0i} + p_e \int a(t_{0i}) f(t_{0i}|k_i=1) dt_{0i}\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[k_i a(t_{0i})] &= \int\int k_i a(t_{0i}) f(t_{0i}, k_i) dk_i dt_{0i} \\ &= \sum_{k_i=0,1}\int k_i a(t_{0i}) f(t_{0i}, k_i) dt_{0i} = \int a(t_{0i}) f(t_{0i}, 1) dt_{0i},\end{aligned}$$

and equivalently we can deduce expressions for $\mathbb{E}[c(t_{0i})]$, $\mathbb{E}[d(t_{0i})]$, $\mathbb{E}[k_i c(t_{0i})]$ and $\mathbb{E}[k_i d(t_{0i})]$. Then, it can be derived that the [4,4]th component of the inverse of

$\mathbb{E}\left[\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right]$ is

$$\frac{(1-p_e)^{-1}\int a(t_{0i})f(t_{0i}|k_i=0)dt_{0i}}{\left[\left(\int a(t_{0i})f(t_{0i}|k_i=0)dt_{0i}\right)\left(\int d(t_{0i})f(t_{0i}|k_i=0)dt_{0i}\right)-\left(\int c(t_{0i})f(t_{0i}|k_i=0)dt_{0i}\right)^2\right]}$$
$$+\frac{p_e^{-1}\int a(t_{0i})f(t_{0i}|k_i=1)dt_{0i}}{\left[\left(\int a(t_{0i})f(t_{0i}|k_i=1)dt_{0i}\right)\left(\int d(t_{0i})f(t_{0i}|k_i=1)dt_{0i}\right)-\left(\int c(t_{0i})f(t_{0i}|k_i=1)dt_{0i}\right)^2\right]}. \tag{A.4}$$

In the paper and in our software, we assumed that t_{0i} is normally distributed. We assumed that t_{0i} has mean zero, which can always be achieved by centering at the mean initial time and implies no loss of generality (Kreft et al., 1995), and that it has variance $V(t_0)$. Additionally, we assumed that the variance of t_{0i} is the same within each exposure group. In Appendix A.1.2 we derived the means of t_{0i} conditional on exposure as

$$\mathbb{E}\left(t_0|k=1\right)=\rho_{\mathrm{e},t_0}\sqrt{\frac{(1-p_e)}{p_e}}\sqrt{V\left(t_0\right)}$$

and

$$\mathbb{E}\left(t_0|k=0\right)=-\rho_{\mathrm{e},t_0}\sqrt{\frac{p_e}{(1-p_e)}}\sqrt{V\left(t_0\right)}.$$

Using results from Appendix A.1.2, we find that

$$\begin{aligned}V\left(t_0|k=1\right)&=\mathbb{E}\left(t_0^2|k=1\right)-\left[\mathbb{E}\left(t_0|k=1\right)\right]^2\\&=\frac{V\left(t_0\right)\left[p_e+\rho_{\mathrm{e},t_0}^2(1-2p_e)\right]}{p_e}-\rho_{\mathrm{e},t_0}^2\frac{(1-p_e)}{p_e}V\left(t_0\right)=V\left(t_0\right)\left(1-\rho_{\mathrm{e},t_0}^2\right)\end{aligned}$$

Therefore,

$$f(t_{0i}|k_i=1)=$$
$$\frac{1}{\sqrt{2\pi V\left(t_0\right)\left(1-\rho_{\mathrm{e},t_0}^2\right)}}\exp\left[\frac{-1}{2V\left(t_0\right)\left(1-\rho_{\mathrm{e},t_0}^2\right)}\left(t_{0i}-\rho_{\mathrm{e},t_0}\sqrt{\frac{(1-p_e)}{p_e}}\sqrt{V\left(t_0\right)}\right)^2\right]$$

and

$$f(t_{0i}|k_i=0)=$$
$$\frac{1}{\sqrt{2\pi V\left(t_0\right)\left(1-\rho_{\mathrm{e},t_0}^2\right)}}\exp\left[\frac{-1}{2V\left(t_0\right)\left(1-\rho_{\mathrm{e},t_0}^2\right)}\left(t_{0i}+\rho_{\mathrm{e},t_0}\sqrt{\frac{p_e}{(1-p_e)}}\sqrt{V\left(t_0\right)}\right)^2\right].$$

Our program uses these distributions to compute (A.4) numerically. For CMD, the procedure is exactly the same but using the matrix

$$\mathbf{W}_i = \begin{pmatrix} 1 & 0 & k_i \\ 0 & 1 & 0 \end{pmatrix}$$

A.9 Proof that r_{opt} is the same for both the cost constraint and the power constraint, and reduces to the solution to the unconstrained problem (4.2), but N_{opt} depends upon the constraint

The power optimization problem is

$$\underset{r}{Max} \;\; \Phi\left[\frac{\sqrt{N}\left|(\mathbf{c}'\mathbf{B})_{H_A}\right|}{\sqrt{\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}}} - z_{1-\alpha/2}\right] \text{subject to } COST = Nc_1 + \frac{Nrc_1}{\kappa}.$$

The cost constraint,

$$N = \frac{\kappa COST}{c_1(\kappa + r)},$$

can be plugged in the optimization function to obtain the unconstrained problem

$$\underset{r}{Max} \;\; \Phi\left[\frac{\sqrt{\frac{\kappa COST}{c_1(\kappa+r)}}\left|(\mathbf{c}'\mathbf{B})_{H_A}\right|}{\sqrt{\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}}} - z_{1-\alpha/2}\right]$$

Since Φ is a monotone function, this is equivalent to

$$\underset{r}{Max} \;\; \frac{\sqrt{\frac{\kappa COST}{c_1(\kappa+r)}}\left|(\mathbf{c}'\mathbf{B})_{H_A}\right|}{\sqrt{\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}}} - z_{1-\alpha/2}.$$

Removing positive constant terms with respect to r, it is equivalent to

$$\underset{r}{Max} \;\; \frac{1}{(\kappa + r)\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}},$$

which is in turn equivalent to $\underset{r}{Min} \; (\kappa + r)\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}$. Once r_{opt} is found solving this minimization problem, N_{opt} would be

$$N_{opt} = \frac{\kappa \, COST}{c_1\,(\kappa + r_{opt})}.$$

The cost optimization problem is

$$\underset{r}{Min}\ \ Nc_1 + \frac{Nrc_1}{\kappa} \text{ subject to } \Phi\left[\frac{\sqrt{N}\left|(\mathbf{c}'\mathbf{B})_{H_A}\right|}{\sqrt{\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r)\mathbf{c}}} - z_{1-\alpha/2}\right] = \pi.$$

Noting that

$$Nc_1 + \frac{Nrc_1}{\kappa} = Nc_1\left(\frac{\kappa + r}{\kappa}\right)$$

and that from the power constraint

$$N = \frac{(\mathbf{c}'\,\mathbf{\Sigma}_{\mathrm{B}}(r)\,\mathbf{c})\left(z_{1-\alpha/2} + z_{\pi}\right)^2}{\left((\mathbf{c}'\mathrm{B})_{H_A}\right)^2},$$

this is equivalent to the unconstrained problem

$$\underset{r}{Min}\ \ \frac{(\mathbf{c}'\,\mathbf{\Sigma}_{\mathrm{B}}(r)\,\mathbf{c})\left(z_{1-\alpha/2} + z_{\pi}\right)^2}{\left((\mathbf{c}'\mathrm{B})_{H_A}\right)^2} c_1\left(\frac{\kappa + r}{\kappa}\right)$$

Removing positive constant terms with respect to r, the problem becomes $\underset{r}{Min}\ \ (\kappa + r)\,(\mathbf{c}'\,\mathbf{\Sigma}_{\mathrm{B}}(r)\,\mathbf{c})$, which is equivalent to the minimization problem obtained before. Thus, given κ, $\mathbf{c}$ and $\mathbf{\Sigma}_{\mathrm{B}}(r)$, the same r_{opt} maximizes power and minimizes cost. For the cost problem, once r_{opt} is found solving the minimization problem,

$$N_{opt} = \frac{(\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}(r_{opt})\mathbf{c})\ \left(z_{\pi} + z_{1-\alpha/2}\right)^2}{(\mathbf{c}'\mathrm{B}_{H_A})^2}.$$

A.10 Derivation of (N_{opt}, r_{opt})

A.10.1 (N_{opt}, r_{opt}) under LDD and fixed s, for CS

The optimal r solves $\underset{r}{Min}\ (\kappa{+}r)\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ (Appendix A.9). Plugging in the appropriate value for $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ from Table 1, the problem under LDD, CS and fixed s is

$$\underset{r}{Min}\ (\kappa + r)\frac{12\sigma^2(1-\rho)}{p_e(1-p_e)s^2 r(r+1)(r+2)}.$$

Removing positive constant terms with respect to r, this problem becomes

$$\underset{r}{Min}\quad F(r) = \frac{(\kappa + r)}{r\ (r+1)(r+2)}.$$

Since

$$\frac{\partial F}{\partial r} = \frac{-2\kappa - 6\kappa r - 3r^2 - 3\kappa r^2 - 2r^3}{r^2(r+1)^2(r+2)^2} < 0 \ \forall \kappa,$$

$F(r)$ decreases as r increases, and $r_{opt} \to \infty$ subject to the cost constraint.

A.10.2 (N_{opt}, r_{opt}) under LDD, RS and fixed s

The optimal r solves $\underset{r}{Min}\ (\kappa+r)\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ (Appendix A.9). Plugging in the appropriate value for $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ from Table 1, the problem under LDD, RS and fixed s is

$$\underset{r}{Min}\ (\kappa+r)\left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2 p_e(1-p_e)}\right)\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right)$$

Removing positive constant terms with respect to r, this problem becomes

$$\underset{r}{Min} \quad G(r) = (\kappa+r)\left(\frac{1}{r(r+1)(r+2)} + \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right)$$

The solution, r_{opt}, solves

$$\begin{aligned}\frac{\partial G}{\partial r} &= \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\left(\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right) + \frac{-2\kappa - 6\kappa r - 3r^2 - 3\kappa r^2 - 2r^3}{r^2(r+1)^2(r+2)^2} \\ &= \left(\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)\left(\frac{1}{\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}\right) + \frac{\partial F}{\partial r} = 0,\end{aligned}$$

where $\frac{\partial F}{\partial r}$ is the derivative of the objective function $F(r)$ for the analogous problem under compound symmetry (Appendix A.10.1). We showed in Appendix A.10.1 that $\frac{\partial F}{\partial r}$ is always negative, and since

$$\frac{\partial^2 F}{\partial r^2} = \frac{2\left(4\kappa + 18\kappa r + 33\kappa r^2 + 7r^3 + 24\kappa r^3 + 9r^4 + 6\kappa r^4 + 3r^5\right)}{r^3\left(r+1\right)^3(r+2)^3} \geqslant 0,$$

$\frac{\partial F}{\partial r}$ is also an increasing function of r. In addition, $\lim_{r\to\infty}\frac{\partial F}{\partial r} = 0^-$. Since $\frac{\partial G}{\partial r}$ is $\frac{\partial F}{\partial r}$ plus a constant, $\frac{\partial G}{\partial r}$ will equal 0 at some interior point of r between 1 and ∞. Since

$$\frac{\partial^2 G}{\partial r^2} = \frac{2\left(4\kappa + 18\kappa r + 33\kappa r^2 + 7r^3 + 24\kappa r^3 + 9r^4 + 6\kappa r^4 + 3r^5\right)}{r^3\left(r+1\right)^3(r+2)^3} \geqslant 0$$

for all $r > 0$, $G(r)$ is convex and the point that solves $\frac{\partial G}{\partial r} = 0$ is a global minimum and therefore it is r_{opt}. Now,

$$\frac{\partial G}{\partial r} = 0 \Leftrightarrow \kappa = \frac{r_{opt}^2\left(-(3+2r_{opt})\tilde{r}(\tilde{r}+1)(\tilde{r}+2) + (r_{opt}+1)^2(r_{opt}+2)^2\frac{\rho_{b_1,s,\tilde{r}}}{1-\rho_{b_1,s,\tilde{r}}}\right)}{\left(2+6r_{opt}+3r_{opt}^2\right)\tilde{r}(\tilde{r}+1)(\tilde{r}+2)}.$$

Figure 8 of the paper shows r_{opt} for several values of κ and $\rho_{b_1,s,\tilde{r}}$.

A.10.3 (N_{opt}, r_{opt}) under LDD, CS and fixed τ

As shown in Appendix A.9, the optimal r solves $\underset{r}{Min}\ (\kappa+r)\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$. Plugging in the appropriate value for $\mathbf{c}'\boldsymbol{\Sigma}_{\mathrm{B}}\mathbf{c}$ from Table 1, the problem under LDD, CS and fixed τ is

$$\underset{r}{Min}\ (\kappa+r)\frac{12\sigma^2(1-\rho)r}{p_e(1-p_e)\tau^2(r+1)(r+2)}.$$

Removing positive constant terms with respect to r, this problem becomes

$$\underset{r}{Min} \quad H(r) = \frac{(\kappa+r)\,r}{(r+1)(r+2)}.$$

Taking derivatives with respect to r, r_{opt} solves

$$\frac{\partial H}{\partial r} = \frac{(3-\kappa)r^2+4r+2\kappa}{(r+1)^2\,(r+2)^2} = 0.$$

For $\kappa < 3$ the derivative is positive. Therefore, when $\kappa < 3$, $H(r)$ increases with r and, consequently, the minimum is at $r = 1$. If $\kappa > 3$, the derivative equals 0 at

$$r = \frac{2 \pm \sqrt{2}\sqrt{2-3\kappa+\kappa^2}}{\kappa-3},$$

which gives a positive solution only at

$$r = \frac{2 + \sqrt{2}\sqrt{2-3\kappa+\kappa^2}}{\kappa-3}.$$

Now, we need to check whether at this point there is a maximum or a minimum of $H(r)$. The second derivative of $H(r)$ is

$$\frac{\partial^2 H}{\partial r^2} = \frac{2\,(4-6\kappa-6\kappa r-6r^2-3r^3+\kappa r^3)}{(r+1)^3(r+2)^3}.$$

We evaluated the second derivative at the point

$$r = \frac{2 + \sqrt{2}\sqrt{2-3\kappa+\kappa^2}}{\kappa-3}$$

with Mathematica (Wolfram Research Inc., 2005) and obtained

$$24 + \frac{3\sqrt{2}}{\sqrt{(\kappa-2)(\kappa-1)}}\Bigg)\Bigg(\kappa - 17\sqrt{2}\sqrt{(\kappa-2)(\kappa-1)} - \frac{7\sqrt{2}}{\sqrt{(\kappa-2)(\kappa-1)}} - 40.$$

This expression can be proven to be negative for all $\kappa > 3$. Therefore, $H(r)$ has a maximum at

$$r = \frac{2 + \sqrt{2}\sqrt{2 - 3\kappa + \kappa^2}}{\kappa - 3},$$

while we were looking for a minimum. Since this is the only local maximum or minimum of $H(r)$, and $H(r)$ is continuous, the global minimum of $H(r)$ will be at $r = 1$ or at $r = \infty$. The global minimum will be at $r = \infty$ if we can find a value of r such that

$$H(r) < H(1) = \frac{(1+\kappa)}{6}.$$

With a little bit of algebra , we get

$$H(r) = \frac{r(r+\kappa)}{(r+1)(r+2)} < \frac{(1+\kappa)}{6} \Leftrightarrow r^2(\kappa - 5) + r(-3\kappa + 3) + 2(\kappa + 1) > 0,$$

which has roots at $r = 1$ and $r = \frac{2(\kappa+1)}{\kappa-5}$. If $\kappa < 5$, then $r = \frac{2(\kappa+1)}{\kappa-5} < 0$, outside of its valid range. The global minimum is then $r_{opt} = 1$,

$$N_{opt} = \frac{\kappa COST}{c_1(\kappa+1)}$$

or

$$N_{opt} = \frac{2\sigma^2(1-\rho)\left(z_\pi + z_{1-\alpha/2}\right)^2}{\tau^2 p_e(1-p_e)\gamma_3^2}$$

for the power maximization or cost minimization problems, respectively. If $\kappa > 5$ then $r^2(\kappa - 5) + r(-3\kappa + 3) + 2(\kappa + 1)$ is a convex function and is greater than 0 for $r > \frac{2(\kappa+1)}{\kappa-5}$. Thus, the global minima will be at $r = \infty$. In practice, when $\kappa > 5$, we will choose r as large as possible provided $r > \frac{2(\kappa+1)}{\kappa-5}$ and then find N_{opt} to satisfy the cost or power constraint. If there is no feasible value of r greater than $\frac{2(\kappa+1)}{\kappa-5}$ then one will choose $r = 1$.

A.10.4 (N_{opt}, r_{opt}) under LDD, RS and fixed τ

The optimal r solves $\underset{r}{Min}\ (\kappa+r)\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ (Appendix A.9). Plugging in the appropriate value for $\mathbf{c}'\mathbf{\Sigma}_{\mathrm{B}}\mathbf{c}$ from Table1, the problem under LDD, RS and fixed τ is

$$\underset{r}{Min}\ (\kappa + r)\left(\frac{12\sigma^2(1-\rho_{t_0})}{s^2 p_e(1-p_e)}\right)\left(\left(\frac{r}{(r+1)(r+2)} + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)\left(\right.\right.$$

Removing positive constant terms with respect to r, this problem becomes

$$\underset{r}{Min} \quad I(r) = (\kappa + r)\left(\frac{r}{(r+1)(r+2)} + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)$$

Taking derivatives with respect to r, r_{opt} solves

$$\frac{\partial I}{\partial r} = \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)} + \frac{(3-\kappa)r^2+4r+2\kappa}{(r+1)^2(r+2)^2} = \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)} + \frac{\partial H}{\partial r} = 0,$$

where $\frac{\partial H}{\partial r}$ is the derivative of the objective function $H(r)$ for the analogous problem under CS, given in Appendix A.10.3. There we showed that if $\kappa < 3$ then $\frac{\partial H}{\partial r}$ was strictly positive for all r, and therefore so is $\frac{\partial I}{\partial r}$. Thus, if $\kappa < 3$, $I(r)$ is minimized at $r_{opt} = 1$. For $\kappa > 3$, we know that $\frac{\partial H}{\partial r}$ is continuous, has only one root in the range of interest and it can be shown that $\lim_{r\to\infty} \frac{\partial H}{\partial r} = 0^-$ and $\frac{\partial H(1)}{\partial r} = \frac{7+\kappa}{36}$. It can also be shown with Mathematica (Wolfram Research Inc., 2005) that $\frac{\partial^2 H}{\partial r^2}$ has only one real root, $r*$. Therefore, $\frac{\partial H}{\partial r}$ is positive at $r = 1$, it crosses 0 at the root

$$r = \frac{2+\sqrt{2}\sqrt{2-3\kappa+\kappa^2}}{\kappa-3},$$

as shown in Appendix A.10.3, it has a minimum at the only root of $\frac{\partial^2 H}{\partial r^2}$ and it increases again towards zero, where it reaches an asymptote. Because of the form of $\frac{\partial I}{\partial r}$, it will have a similar shape, since it is equal to $\frac{\partial H}{\partial r}$ but moved upwards by a factor of

$$\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}.$$

Therefore, $\frac{\partial I}{\partial r}$ will have zero roots if

$$\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)} > \frac{\partial H(r*)}{\partial r},$$

or two roots otherwise. In the first case, when $\frac{\partial I}{\partial r}$ has zero roots, $\frac{\partial I}{\partial r}$ is always positive and therefore $I(r)$ increases as r increases and the minimum of $I(r)$ is at $r_{opt} = 1$. In the second case, $\frac{\partial I}{\partial r}$ has two roots, which solve

$$\kappa = \frac{r(4+3r)(\tilde{r}+1)(\tilde{r}+2) + \tilde{r}(r+1)^2(r+2)^2\left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1-\rho_{b_1,\tau,\tilde{r}}}\right)}{(r^2-2)(\tilde{r}+1)(\tilde{r}+2)}.$$

Also, $\frac{\partial^2 H}{\partial r^2} = \frac{\partial^2 I}{\partial r^2}$, and $\frac{\partial^2 H}{\partial r^2}$ is continuous and it has only one root at $r*$. $\frac{\partial^2 H}{\partial r^2}$ is negative for $r < r*$ and positive for $r > r*$. Since $r*$ lies between the first and second roots of $\frac{\partial I}{\partial r}$, it can be concluded that the first root is a maximum of $I(r)$ and the second root is a minimum of $I(r)$. The function $I(r)$ has, therefore, two local minima, one at $r = 1$ and the other at the second root of $\frac{\partial I}{\partial r}$. To find out when the second root is the global minimum of $I(r)$ we need to solve $I(1) > I(r)$, where

$$I(1) = (\kappa + 1)\left(\frac{1}{6} + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1 - \rho_{b_1,\tau,\tilde{r}}}\right)\left(\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)\right)$$

and

$$I(r) = (\kappa + r)\left(\left(\frac{r}{(r+1)(r+2)}\right) + \left(\frac{\rho_{b_1,\tau,\tilde{r}}}{1 - \rho_{b_1,\tau,\tilde{r}}}\right)\frac{\tilde{r}}{(\tilde{r}+1)(\tilde{r}+2)}\right)$$

Provided $r > 2$, this is equivalent to

$$\rho_{b_1,\tau,\tilde{r}} < \frac{\left[-2(\kappa+1) + (\kappa-5)r\right](\tilde{r}+1)(\tilde{r}+2)}{6\tilde{r}(r+1)(r+2) + \left[-2(\kappa+1) + (\kappa-5)r\right](\tilde{r}+1)(\tilde{r}+2)}.$$

The condition only makes sense if $-2(\kappa+1) + (\kappa-5)r > 0$, which is equivalent to the conditions $\kappa > 5$ and $r > \frac{2(\kappa-1)}{\kappa-5}$. Figure 13 of the paper shows this region for different values of $\kappa$and $\rho_{b_1,\tau,\tilde{r}}$, together with a line for the optimal value.

A.11 Demonstration of the use of program `optitxs`

This is the input and output from the program `optitxs` for the calculation of the optimal combination of (N, r) that minimizes the total cost of the study subject to achieving a fixed power under LDD and RS. It is motivated by data from the study examined in section 5. For other examples and a detailed user's guide with many illustrative examples, go to `http://www.hsph.harvard.edu/faculty/spiegelman/optitxs.html`.

```
> long.opt()

* By just pressing <Enter> after each question, the default value,
  shown between square brackets, will be entered.
```

```
* Press <Esc> to quit

Do you want to maximize power subject to a given cost (1) or to
  minimize the total cost subject to a given power (2)[1]? 2

Enter the desired power (0<Pi<1) [0.8]: .8

Are you assuming the time between measurements (s) is fixed (1),
  or the total duration of follow-up (tau) is fixed (2) [1]? 2

Enter the time of follow-up (tau) [1]: 18

Enter the exposure prevalence (pe) (0<=pe<=1) [0.5]: .79

Enter the variance of the time variable at baseline, V(t0)
   (enter 0 if all participants begin at the same time) [0]: 100

Enter the correlation between the time variable at baseline and
 exposure, rho_{e,t0} [0]: 0

Constant mean difference (1) or Linearly divergent difference (2)
 [1]: 2

Will you specify the alternative hypothesis on the absolute
 (beta coefficient) scale (1) or the relative (percent) scale (2)
 [1]? 2

Enter mean response at baseline among unexposed (mu00) [10]: 3.5

Enter the percent change from baseline to end of follow-up among
 unexposed (p2) (e.g. enter 0.10 for a 10% change) [0.1]: -.182

Enter the percent difference between the change from baseline to
 end of follow-up in the exposed group and the unexposed group
 (p3) (e.g. enter 0.10 for a 10% difference) [0.1]: .1

Which covariance matrix are you assuming: compound symmetry (1),
 damped exponential (2) or random slopes (3) [1]? 3

Enter (1) for standard notation (variance of residuals and random
  effects) or (2) for "reliability" notation [1]: 2

Enter the variance of the response given the assumed model
 covariates at baseline (sigma2) [1]: .34
```

```
Enter the reliability coefficient at baseline (0<rho_t0<1)
 [0.8]: .877

Enter the trial value of the number of measurements at which the
        slope reliability will be provided (\tilde r>0 ) [5]: 6

Enter the slope reliability for 6 repeated measurements
 (0<rho_{b1,s,\tilde r}<1  or 0<rho_{b1,tau,\tilde r}<1)
 [0.1]: .364

Enter the correlation between the random effects of slope
 and intercept (-1<rho[b0,b1]<1) [0]: -.32

Enter the cost of the first observation of each subject (c1>0)
 [80]: 80

Enter the ratio of costs between the first measure and the rest
 (kappa) [2]: 20

Cost optimization problem (min cost for a given power):
   Optimal r= 12 , Optimal N= 732 , Power= 0.8 ,Cost= 93696

Slope reliability at r= 12 :  0.4818737
```

References

L. Frison and S. J. Pocock. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Stat Med*, 11(13):1685–704, 1992.

L. J. Frison and S. J. Pocock. Linearly divergent treatment effects in clinical trials with repeated measures: efficient analysis using summary statistics. *Stat Med*, 16(24):2855–72, 1997.

Franklin A. Graybill. *Matrices with applications in statistics*. Wadsworth, Belmont, Calif., 2nd edition, 1983.

I. G. G. Kreft, J. Leeuw, and L. S. Aiken. The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, 30(1):1–21, 1995.

John M. Lachin. *Biostatistical methods: the assessment of relative risks*. Wiley series in probability and statistics. Wiley, New York, 2000.

S. Puntanen and G. P. H. Styan. The equality of the ordinary least squares estimator and the best linear unbiased estimator. *The American Statistician*, 43(3):153–61, 1989.

Neil H. Timm. *Applied multivariate analysis*. Springer texts in statistics. Springer, New York, 2002.

Wolfram Research Inc. *Mathematica, version 5.2*. Champaign, Illinois, 2005.