Web-based Supporting Materials for

“Optimal combination of number of participants and number of repeated measurements in longitudinal studies with time-varying exposure”

by J. Barrera-Gómez, D. Spiegelman and X. Basagaña.

Contents

A Derivation of $\tilde{\sigma}^2$ for basic scenarios

A.1 Time-invariant exposure

A.1.1 CMD

We assume a CS(σ, ρ) response covariance structure, no missing data and constant exposure prevalence (i.e., $p_{ej} = p_e,\ \forall j = 0, \ldots, r$).

From model (1) we have

$$\begin{aligned}
\mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i &= \begin{pmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 \\ 0 & s & \cdots & sk & \cdots & sr \\ E_i & E_i & \cdots & E_i & \cdots & E_i \end{pmatrix} \begin{pmatrix} \nu_{00} & \cdots & \nu_{0r} \\ \vdots & \ddots & \vdots \\ \nu_{r0} & \cdots & \nu_{rr} \end{pmatrix} \begin{pmatrix} 1 & 0 & E_i \\ 1 & s & E_i \\ \vdots & \vdots & \vdots \\ 1 & sj & E_i \\ \vdots & \vdots & \vdots \\ 1 & sr & E_i \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 \\ 0 & s & \cdots & sk & \cdots & sr \\ E_i & E_i & \cdots & E_i & \cdots & E_i \end{pmatrix} \begin{pmatrix} \sum_{j=0}^r \nu_{0j} & s\sum_{j=0}^r j\nu_{0j} & E_i \sum_{j=0}^r \nu_{0j} \\ \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{mj} & s\sum_{j=0}^r j\nu_{mj} & E_i \sum_{j=0}^r \nu_{mj} \\ \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{rj} & s\sum_{j=0}^r j\nu_{rj} & E_i \sum_{j=0}^r \nu_{rj} \end{pmatrix} \\
&= \begin{pmatrix} \omega_{00} & s\omega_{10} & E_i\omega_{00} \\ s\omega_{10} & s^2\omega_{11} & sE_i\omega_{10} \\ E_i\omega_{00} & sE_i\omega_{10} & E_i^2\omega_{00} \end{pmatrix},
\end{aligned}$$

where $s = \frac{1}{r}$ is the elapsed time between two consecutives measurements in units of the fixed total follow-up time and

$$\omega_{pq} := \sum_{j=0}^{r} \sum_{k=0}^{r} j^p k^q \nu_{jk}.$$

Since $E_i \sim \text{Bernoulli}(p_e)$, we have

$$\mathbb{E}_X \left[\mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right] = \begin{pmatrix} \omega_{00} & s\omega_{10} & p_e\omega_{00} \\ s\omega_{10} & s^2\omega_{11} & sp_e\omega_{10} \\ p_e\omega_{00} & sp_e\omega_{10} & p_e\omega_{00} \end{pmatrix}$$

and

$$\det(\mathbb{E}_X \left[\mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right]) = \omega_{00}(\omega_{00}\omega_{11} - \omega_{10}^2) s^2 p_e (1 - p_e).$$

We are interested in the [3,3]–th element of $(\mathbb{E}_X [\mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i])^{-1}$ which is

$$\frac{\begin{vmatrix} \omega_{00} & s\omega_{10} \\ s\omega_{10} & s^2\omega_{11} \end{vmatrix}}{\det(\mathbb{E}_X [\mathbf{X}_i' \boldsymbol{\Sigma}^{-1} \mathbf{X}_i])} = \frac{1}{\omega_{00} p_e (1 - p_e)}.$$

Thus,

$$\tilde{\sigma}^2 = \frac{1}{p_e(1-p_e)\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}}.$$

If the response correlation structure is CS(σ, ρ), then

$$\mathbf{\Sigma}[j,k] = \begin{cases} \sigma^2 & , \quad j = k \\ \sigma^2\rho & , \quad j \neq k \end{cases}$$

and

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

so

$$\omega_{00} = \sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk} = \frac{r+1}{\sigma^2(\rho r+1)}$$

and then

$$\tilde{\sigma}^2 = \frac{\sigma^2(\rho r+1)}{p_e(1-p_e)(r+1)}.$$

A.1.2 LDD

We assume a CS(σ, ρ) response covariance structure, no missing data and constant exposure prevalence (i.e., $p_{ej} = p_e, \ \forall j = 0, \ldots, r$).

From model (2), we have

$$
\begin{aligned}
\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i &= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_i & E_i & \cdots & E_i \\ 0 & sE_i & \cdots & srE_i \end{pmatrix} \begin{pmatrix} \nu_{00} & \cdots & \nu_{0r} \\ \vdots & \ddots & \vdots \\ \nu_{r0} & \cdots & \nu_{rr} \end{pmatrix} \begin{pmatrix} 1 & 0 & E_i & 0 \\ 1 & s & E_i & sE_i \\ \vdots & \vdots & \vdots & \vdots \\ 1 & sj & E_i & sjE_i \\ \vdots & \vdots & \vdots & \vdots \\ 1 & sr & E_i & srE_i \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_i & E_i & \cdots & E_i \\ 0 & sE_i & \cdots & srE_i \end{pmatrix} . \\
&\quad \begin{pmatrix} \sum_{j=0}^r \nu_{0j} & s\sum_{j=0}^r j\nu_{0j} & E_i\sum_{j=0}^r \nu_{0j} & sE_i\sum_{j=0}^r j\nu_{0j} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{mj} & s\sum_{j=0}^r j\nu_{mj} & E_i\sum_{j=0}^r \nu_{mj} & sE_i\sum_{j=0}^r j\nu_{mj} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{rj} & s\sum_{j=0}^r j\nu_{rj} & E_i\sum_{j=0}^r \nu_{rj} & sE_i\sum_{j=0}^r j\nu_{rj} \end{pmatrix} \\
&= \begin{pmatrix} \omega_{00} & s\omega_{10} & E_i\omega_{00} & sE_i\omega_{10} \\ s\omega_{10} & s^2\omega_{11} & sE_i\omega_{10} & s^2E_i\omega_{11} \\ E_i\omega_{00} & sE_i\omega_{10} & E_i^2\omega_{00} & sE_i^2\omega_{10} \\ sE_i\omega_{10} & s^2E_i\omega_{11} & sE_i^2\omega_{10} & s^2E_i^2\omega_{11} \end{pmatrix} .
\end{aligned}
$$

Then,

$$
\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right] = \begin{pmatrix} \omega_{00} & s\omega_{10} & p_e\omega_{00} & sp_e\omega_{10} \\ s\omega_{10} & s^2\omega_{11} & sp_e\omega_{10} & s^2p_e\omega_{11} \\ p_e\omega_{00} & sp_e\omega_{10} & p_e\omega_{00} & sp_e\omega_{10} \\ sp_e\omega_{10} & s^2p_e\omega_{11} & sp_e\omega_{10} & s^2p_e\omega_{11} \end{pmatrix} .
$$

and

$$
\det(\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right]) = \left[(\omega_{00}\omega_{11} - \omega_{10}^2)s^2p_e(1-p_e)\right]^2
$$

We are interested in the [4,4]–th element of $(\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right])^{-1}$ which is

$$
\frac{\begin{vmatrix} \omega_{00} & s\omega_{10} & p_e\omega_{00} \\ s\omega_{10} & s^2\omega_{11} & sp_e\omega_{10} \\ p_e\omega_{00} & sp_e\omega_{10} & p_e\omega_{00} \end{vmatrix}}{\det(\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right])} = \frac{\omega_{00}}{(\omega_{00}\omega_{11} - \omega_{10}^2)s^2p_e(1-p_e)} .
$$

Thus,

$$
\tilde{\sigma}^2 = \frac{r^2\sum_{j=0}^r\sum_{k=0}^r \nu_{jk}}{p_e(1-p_e)\left[\left(\sum_{j=0}^r\sum_{k=0}^r \nu_{jk}\right)\left(\sum_{j=0}^r\sum_{k=0}^r jk\nu_{jk}\right) - \left(\sum_{j=0}^r\sum_{k=0}^r j\nu_{jk}\right)^2\right]} .
$$

If the response covariance structure is CS(σ, ρ), then

$$\omega_{10} = \sum_{j=0}^{r}\sum_{k=0}^{r} j\nu_{jk} = \frac{r(r+1)}{2\sigma^2(\rho r+1)}$$

and

$$\omega_{11} = \sum_{j=0}^{r}\sum_{k=0}^{r} jk\nu_{jk} = \frac{r(r+1)(\rho r^2+(4-\rho)r+2)}{12\sigma^2(1-\rho)(\rho r+1)}$$

and then

$$\tilde{\sigma}^2 = \frac{12\sigma^2(1-\rho)r}{p_e(1-p_e)(r+1)(r+2)}.$$

A.2 Time-varying exposure

A.2.1 CMD

We assume a CS(σ, ρ) response covariance structure, no missing data and constant prevalence of the exposure ($p_{ej} = p_e,\ \forall j = 0, \ldots, r$). We consider a general exposure covariance matrix with elements

$$\mathbf{\Sigma}_E[j,k] = \begin{cases} \mathrm{Var}(E_{ij}) = p_e(1-p_e) & , \quad j = k \\ \sigma_{e_{jk}} := \mathrm{Cov}(E_{ij}, E_{ik}) = \rho_{e_{jk}} p_e(1-p_e) & , \quad j \neq k \end{cases}$$

where $\rho_{e_{jk}} = \mathrm{Cor}(E_{ij}, E_{ik})$ is correlation between the j–th and the k–th exposure measurements, assumed to be common to all participants.

From model (1), we have

$$\begin{aligned}
\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i &= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_{i0} & E_{i1} & \cdots & E_{ir} \end{pmatrix} \begin{pmatrix} \nu_{00} & \cdots & \nu_{0r} \\ \vdots & \ddots & \vdots \\ \nu_{r0} & \cdots & \nu_{rr} \end{pmatrix} \begin{pmatrix} 1 & 0 & E_{i0} \\ 1 & s & E_{i1} \\ \vdots & \vdots & \vdots \\ 1 & sj & E_{ij} \\ \vdots & \vdots & \vdots \\ 1 & sr & E_{ir} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_{i0} & E_{i1} & \cdots & E_{ir} \end{pmatrix} \begin{pmatrix} \sum_{j=0}^r \nu_{0j} & s\sum_{j=0}^r j\nu_{0j} & \sum_{j=0}^r \nu_{0j}E_{ij} \\ \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{mj} & s\sum_{j=0}^r j\nu_{mj} & \sum_{j=0}^r \nu_{mj}E_{ij} \\ \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{rj} & s\sum_{j=0}^r j\nu_{rj} & \sum_{j=0}^r \nu_{rj}E_{ij} \end{pmatrix} \\
&= \begin{pmatrix} \omega_{00} & s\omega_{10} & \phi_0 \\ s\omega_{10} & s^2\omega_{11} & s\phi_1 \\ \phi_0 & s\phi_1 & \epsilon \end{pmatrix},
\end{aligned}$$

where

$$\phi_q := \sum_{j=0}^r \sum_{k=0}^r k^q \nu_{jk} E_{ij},\ q = 0, 1; \qquad \epsilon := \sum_{j=0}^r \sum_{k=0}^r \nu_{jk} E_{ij} E_{ik}.$$

Now we compute

$$\begin{aligned}
\mathbb{E}_X[\phi_q] &= \mathbb{E}_X\left[\sum_{j=0}^r \sum_{k=0}^r k^q \nu_{jk} E_{ij}\right] = \sum_{j=0}^r \sum_{k=0}^r k^q \nu_{jk} \mathbb{E}_X[E_{ij}] \\
&= p_e \sum_{j=0}^r \sum_{k=0}^r k^q \nu_{jk} = p_e \omega_{q0}, \quad q = 0, 1.
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}_X\left[\epsilon\right] &= \mathbb{E}_X\left[\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}E_{ij}E_{ik}\right] = \sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\mathbb{E}_X\left[E_{ij}E_{ik}\right] \\
&= \sum_{j=0}^{r}\nu_{jj}\mathbb{E}_X\left[E_{ij}^2\right] + \sum_{j=0,\,j\neq k}^{r}\sum_{k=0}^{r}\nu_{jk}\mathbb{E}_X\left[E_{ij}E_{ik}\right] \\
&= \sum_{j=0}^{r}\nu_{jj}\mathbb{E}_X[E_{ij}] + \sum_{j=0,\,j\neq k}^{r}\sum_{k=0}^{r}\nu_{jk}\left\{\mathrm{Cov}(E_{ij},E_{ik}) + \mathbb{E}_X\left[E_{ij}\right]\mathbb{E}_X\left[E_{ik}\right]\right\} \\
&= p_e\sum_{j=0}^{r}\nu_{jj} + \sum_{j=0,\,j\neq k}^{r}\sum_{k=0}^{r}\nu_{jk}\left[\rho_{e_{jk}}p_e(1-p_e)+p_e^2\right] \\
&= p_e\sum_{j=0}^{r}\nu_{jj} + p_e(1-p_e)\sum_{j=0,\,j\neq k}^{r}\sum_{k=0}^{r}\nu_{jk}\rho_{e_{jk}} + p_e^2\sum_{j=0,\,j\neq k}^{r}\sum_{k=0}^{r}\nu_{jk} \\
&= p_e\sum_{j=0}^{r}\nu_{jj} + p_e(1-p_e)\left[\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\rho_{e_{jk}} - \sum_{j=0}^{r}\nu_{jj}\rho_{e_{jj}}\right] + p_e^2\left[\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk} - \sum_{j=0}^{r}\nu_{jj}\right] \\
&= p_e^2\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk} + p_e(1-p_e)\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\rho_{e_{jk}} = p_e^2\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk} + \sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\sigma_{e_{jk}}^2 \\
&= p_e^2\,\omega_{00} + \alpha_0,
\end{aligned}
$$

where

$$\alpha_0 := \sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\sigma_{e_{jk}}^2$$

and $\sigma_{e_{jk}}^2$ is the $[j,k]$–th element of the covariance matrix of exposure $\mathbf{\Sigma}_E$.

Then,

$$\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right] = \begin{pmatrix} \omega_{00} & s\omega_{10} & p_e\omega_{00} \\ s\omega_{10} & s^2\omega_{11} & sp_e\omega_{10} \\ p_e\omega_{00} & sp_e\omega_{10} & p_e^2\,\omega_{00}+\alpha_0 \end{pmatrix}$$

and

$$\det(\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right]) = (\omega_{00}\omega_{11}-\omega_{10}^2)s^2\alpha_0\,.$$

We are interested in the [3,3]–th element of $\left(\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right]\right)^{-1}$ which is

$$\frac{\begin{vmatrix} \omega_{00} & s\omega_{10} \\ s\omega_{10} & s^2\omega_{11} \end{vmatrix}}{\det(\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right])} = \frac{1}{\alpha_0}$$

and thus

$$\tilde{\sigma}^2 = \frac{1}{\sum_{j=0}^{r}\sum_{k=0}^{r}\nu_{jk}\sigma_{e_{jk}}^2}\,.$$

If the response covariance structure is CS(σ, ρ), then

$$\nu_{jk} = \begin{cases} \frac{\rho r+1-\rho}{\sigma^2(1-\rho)(\rho r+1)} & , \quad j = k \\ \frac{-\rho}{\sigma^2(1-\rho)(\rho r+1)} & , \quad j \neq k \end{cases}$$

and

$$\begin{aligned}
\alpha_0 &= \sum_{j=0}^{r}\sum_{k=0}^{r} \nu_{jk}\sigma^2_{e_{jk}} = \sum_{j=0}^{r} \nu_{jj}\sigma^2_{e_{jj}} + \sum_{j=0,\, j\neq k}^{r}\sum_{k=0}^{r} \nu_{jk}\sigma^2_{e_{jk}} \\
&= \frac{1}{\sigma^2(1-\rho)(\rho r+1)}\left[p_e(1-p_e)\sum_{j=0}^{r}(\rho r+1-\rho) - \sum_{j=0,\, j\neq k}^{r}\sum_{k=0}^{r}\rho\sigma^2_{e_{jk}}\right] \\
&= \frac{1}{\sigma^2(1-\rho)(\rho r+1)}\left[p_e(1-p_e)(r+1)(\rho r+1-\rho) - \rho\sum_{j=0,\, j\neq k}^{r}\sum_{k=0}^{r}\sigma^2_{e_{jk}}\right] \\
&= \frac{1}{\sigma^2(1-\rho)(\rho r+1)}\left\{p_e(1-p_e)(r+1)(\rho r+1-\rho) - \rho\left[\sum_{j=0}^{r}\sum_{k=0}^{r}\sigma^2_{e_{jk}} - \sum_{j=0}^{r}\sigma^2_{e_{jj}}\right]\right\} \\
&= \frac{1}{\sigma^2(1-\rho)(\rho r+1)}\left\{p_e(1-p_e)(r+1)(\rho r+1-\rho) - \rho\left[\mathrm{sum}(\boldsymbol{\Sigma}_E) - (r+1)p_e(1-p_e)\right]\right\} \\
&= \frac{p_e(1-p_e)(r+1)(\rho r+1) - \rho\,\mathrm{sum}(\boldsymbol{\Sigma}_E)}{\sigma^2(1-\rho)(\rho r+1)} = \frac{p_e(1-p_e)(r+1)[\rho(1-\rho_e)r+1-\rho]}{\sigma^2(1-\rho)(\rho r+1)}
\end{aligned}$$

and then

$$\tilde{\sigma}^2 = \frac{\sigma^2(1-\rho)(\rho r+1)}{p_e(1-p_e)(r+1)[\rho(1-\rho_e)r+1-\rho]}\,,$$

where the intraclass correlation of the exposure is

$$\rho_e = \frac{\mathrm{sum}(\boldsymbol{\Sigma}_E) - \mathrm{Tr}(\boldsymbol{\Sigma}_E)}{r\mathrm{Tr}(\boldsymbol{\Sigma}_E)}\,,$$

where sum() and Tr() denote the sum of the elements and the trace of a matrix respectively.

A.2.2 LDD

We assume a CS(σ, ρ) response covariance structure, no missing data and constant exposure prevalence (i.e., $p_{ej} = p_e,\ \forall j = 0, \ldots, r$). We consider a CS structure for exposure covariance matrix:

$$\mathbf{\Sigma}_E[j,k] = \begin{cases} \mathrm{Var}(E_{ij}) = p_e(1-p_e) & ,\quad j = k \\ \sigma_{e_{jk}} = \mathrm{Cov}(E_{ij}, E_{ik}) = \rho_e p_e(1-p_e) & ,\quad j \neq k \end{cases}$$

where $\rho_e = \mathrm{Cor}(E_{ij}, E_{ik})$ is the common correlation of the exposure, also the intraclass correlation of exposure.

From model (2), we have

$$\begin{aligned}
\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i &= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_{i0} & E_{i0} & \cdots & E_{i0} \\ 0 & E_{i1}^\star & \cdots & E_{ir}^\star \end{pmatrix} \begin{pmatrix} \nu_{00} & \cdots & \nu_{0r} \\ \vdots & \ddots & \vdots \\ \nu_{r0} & \cdots & \nu_{rr} \end{pmatrix} \begin{pmatrix} 1 & 0 & E_{i0} & 0 \\ 1 & s & E_{i0} & E_{i1}^\star \\ \vdots & \vdots & \vdots & \vdots \\ 1 & sr & E_{i0} & E_{ir}^\star \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & s & \cdots & sr \\ E_{i0} & E_{i0} & \cdots & E_{i0} \\ 0 & E_{i1}^\star & \cdots & E_{ir}^\star \end{pmatrix} \cdot \\
&\quad \begin{pmatrix} \sum_{j=0}^r \nu_{0j} & s\sum_{j=0}^r j\nu_{0j} & E_{i0}\sum_{j=0}^r \nu_{0j} & \sum_{j=0}^r \nu_{0j}E_{ij}^\star \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{mj} & s\sum_{j=0}^r j\nu_{mj} & E_{i0}\sum_{j=0}^r \nu_{mj} & \sum_{j=0}^r \nu_{mj}E_{ij}^\star \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=0}^r \nu_{rj} & s\sum_{j=0}^r j\nu_{rj} & E_{i0}\sum_{j=0}^r \nu_{rj} & \sum_{j=0}^r \nu_{rj}E_{ij}^\star \end{pmatrix} \\
&= \begin{pmatrix} \omega_{00} & s\omega_{10} & E_{i0}\omega_{00} & \eta_1 \\ s\omega_{10} & s^2\omega_{11} & sE_{i0}\omega_{10} & s\eta_2 \\ E_{i0}\omega_{00} & sE_{i0}\omega_{10} & E_{i0}^2\omega_{00} & E_{i0}\eta_1 \\ \eta_1 & s\eta_2 & E_{i0}\eta_1 & \eta_3 \end{pmatrix}
\end{aligned}$$

where

$$\eta_1 := \sum_{k=0}^r \sum_{j=1}^r \nu_{jk}E_{ij}^\star, \quad \eta_2 := \sum_{k=1}^r \sum_{j=1}^r k\nu_{jk}E_{ij}^\star, \quad \eta_3 := \sum_{k=1}^r \sum_{j=1}^r \nu_{jk}E_{ij}^\star E_{ik}^\star .$$

Now, for $j \geqslant 1$,

$$\mathbb{E}_X\left[E_{ij}^\star\right] = \mathbb{E}_X\left[s\sum_{m=1}^j E_{im}\right] = s\sum_{m=1}^j \mathbb{E}_X\left[E_{im}\right] = p_e sj\,,$$

$$\begin{aligned}
\mathbb{E}_X\left[{E_{ij}^\star}^2\right] &= s^2\mathbb{E}_X\left[\left(\sum_{m=1}^j E_{im}\right)^2\right] = s^2\mathbb{E}_X\left[\sum_{l=1}^j\sum_{m=1}^j E_{im}E_{il}\right] \\
&= s^2\left(j\mathbb{E}_X\left[E_{ij}^2\right] + (j^2-j)\mathbb{E}_X\left[E_{ij}E_{ik\neq j}\right]\right) \\
&= s^2 j\left\{\mathbb{E}_X\left[E_{ij}\right] + (j-1)\left[\rho_e \mathrm{Var}(E_{ij}) + \left(\mathbb{E}_X\left[E_{ij}\right]\right)^2\right]\right\} \\
&= p_e s^2 j\left[1 + (j-1)\lambda\right],
\end{aligned}$$

where

$$\lambda := \rho_e(1-p_e)+p_e,$$

and

$$\begin{aligned}
\mathbb{E}_X\left[E^\star_{ij}E^\star_{ik\neq j}\right] &= s^2\mathbb{E}_X\left[\left(\sum_{l=1}^{\min(j,k)}E_{il}\right)\left(\sum_{m=1}^{\max(j,k)}E_{im}\right)\right]\\
&= s^2\mathbb{E}_X\left[\left(\sum_{l=1}^{\min(j,k)}E_{il}\right)^2+\left(\sum_{l=1}^{\min(j,k)}E_{il}\right)\left(\sum_{m=1+\min(j,k)}^{\max(j,k)}E_{im}\right)\right]\\
&= \mathbb{E}_X\left[E^{\star 2}_{i\ \min(j,k)}\right]+s^2|j-k|\min(j,k)\mathbb{E}_X\left[E_{ij}E_{ik\neq j}\right]\\
&= p_e s^2\min(j,k)\left[1+(\max(j,k)-1)\,\lambda\right].
\end{aligned}$$

Then,

$$\mathbb{E}_X[\eta_1]=\sum_{k=0}^{r}\sum_{j=1}^{r}\nu_{jk}\mathbb{E}_X\left[E^\star_{ij}\right]=p_e s\sum_{k=0}^{r}\sum_{j=1}^{r}j\nu_{jk}=p_e s\,\omega_{10}\,,$$

$$\mathbb{E}_X[\eta_2]=\sum_{k=1}^{r}\sum_{j=1}^{r}k\nu_{jk}\mathbb{E}_X\left[E^\star_{ij}\right]=p_e s\sum_{k=1}^{r}\sum_{j=1}^{r}jk\nu_{jk}=p_e s\,\omega_{11}$$

and

$$\mathbb{E}_X[\eta_3]=\sum_{k=1}^{r}\sum_{j=1}^{r}\nu_{jk}\mathbb{E}_X\left[E^\star_{ij}E^\star_{ik}\right]=p_e s^2[(1-\lambda)\psi_1+\lambda\omega_{11}],$$

where

$$\psi_1 := \sum_{k=1}^{r}\sum_{j=1}^{r}\nu_{jk}\min(j,k).$$

Also,

$$\begin{aligned}
\mathbb{E}_X[E_{i0}\eta_1] &= \sum_{k=0}^{r}\sum_{j=1}^{r}\nu_{jk}\mathbb{E}_X\left[E_{i0}E^\star_{ij}\right]=s\sum_{k=0}^{r}\sum_{j=1}^{r}\nu_{jk}\mathbb{E}_X\left[E_{i0}\sum_{m=1}^{j}E_{im}\right]\\
&= s\sum_{k=0}^{r}\sum_{j=1}^{r}\nu_{jk}\sum_{m=1}^{j}\mathbb{E}_X[E_{i0}E_{im}]=s\sum_{k=0}^{r}\sum_{j=1}^{r}j\nu_{jk}\mathbb{E}_X[E_{i0}E_{im\neq 0}]\\
&= p_e s\lambda\omega_{10}\,.
\end{aligned}$$

Then

$$\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right]=\begin{pmatrix}
\omega_{00} & s\omega_{10} & p_e\omega_{00} & p_e s\omega_{10}\\
s\omega_{10} & s^2\omega_{11} & p_e s\omega_{10} & p_e s^2\omega_{11}\\
p_e\omega_{00} & p_e s\omega_{10} & p_e\omega_{00} & p_e s\lambda\omega_{10}\\
p_e s\omega_{10} & p_e s^2\omega_{11} & p_e s\lambda\omega_{10} & p_e s^2[(1-\lambda)\psi_1+\lambda\omega_{11}]
\end{pmatrix}$$

and

$$\det(\mathbb{E}_X\left[\mathbf{X}_i'\mathbf{\Sigma}^{-1}\mathbf{X}_i\right])=s^4p_e^2(1-p_e)^2(\omega_{00}\omega_{11}-\omega_{10}^2)\left\{\omega_{00}\left[(1-\rho_e)\psi_1+\rho_e\omega_{11}\right]-\rho_e^2\omega_{10}^2\right\}.$$

We are interested in the [4,4]–th element of $\left(\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right]\right)^{-1}$ which is

$$\frac{\begin{vmatrix} \omega_{00} & s\omega_{10} & p_e\omega_{00} \\ s\omega_{10} & s^2\omega_{11} & p_e s\omega_{10} \\ p_e\omega_{00} & p_e s\omega_{10} & p_e\omega_{00} \end{vmatrix}}{\det(\mathbb{E}_X\left[\mathbf{X}_i'\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right])} = \frac{\omega_{00}}{s^2 p_e(1-p_e)\left\{\omega_{00}\left[(1-\rho_e)\psi_1+\rho_e\omega_{11}\right]-\rho_e^2\omega_{10}^2\right\}}$$

and thus

$$\tilde{\sigma}^2 = \frac{r^2\omega_{00}}{p_e(1-p_e)\left\{\omega_{00}\left[(1-\rho_e)\psi_1+\rho_e\omega_{11}\right]-\rho_e^2\omega_{10}^2\right\}}\,.$$

If the response covariance structure is CS(σ, ρ), then

$$\begin{aligned}\psi_1 &= \sum_{k=1}^{r}\sum_{j=1}^{r}\nu_{jk}\min(j,k) = \frac{1}{\sigma^2(1-\rho)(\rho r+1)}\left[(\rho r+1-\rho)\sum_{j=1}^{r} j - 2\rho\sum_{k=1}^{r-1}\sum_{j=1}^{m} j\right] \\ &= \frac{r(r+1)(\rho r+3-\rho)}{6\sigma^2(1-\rho)(\rho r+1)}\end{aligned}$$

and

$$\tilde{\sigma}^2 = \frac{12\sigma^2(1-\rho)(\rho r+1)r}{p_e(1-p_e)(r+1)\left\{\rho\rho_e r^2+\left[2\rho+\rho_e+3(1-\rho)\rho_e(1-\rho_e)\right]r+2\left[1+(1-\rho_e)(2-\rho)\right]\right\}}\,.$$

B Adapting the cost function for dropout

We considered monotone dropout (3), i.e., that losing one individual measurement implies losing all the subsequent measurements of that individual. We assumed that there is no missing data at the first measurement and that each subject that has not dropped out of the study at a given measurement time had a probability π_m of dropout at the subsequent measurement. Thus, the probability and cost of each of the $r+1$ dropout pattern is

# of measurements before the missing	**Probability of the pattern**	**Total cost**
1	π_m	c_1
2	$(1-\pi_m)\pi_m$	$c_1\left(1+\frac{1}{\kappa}\right)$
3	$(1-\pi_m)^2\pi_m$	$c_1\left(1+\frac{2}{\kappa}\right)$
$\vdots$	$\vdots$	$\vdots$
j	$(1-\pi_m)^{j-1}\pi_m$	$c_1\left(1+\frac{j-1}{\kappa}\right)$
$\vdots$	$\vdots$	$\vdots$
r	$(1-\pi_m)^{r-1}\pi_m$	$c_1\left(1+\frac{r-1}{\kappa}\right)$
$r+1$	$(1-\pi_m)^r$	$c_1\left(1+\frac{r}{\kappa}\right)$

and then the expected value of the total cost for N participants is

$$\begin{aligned}
\mathbb{E}(\text{Cost}) &= Nc_1\left[\pi_m+\pi_m\sum_{j=1}^{r-1}(1-\pi_m)^j+(1-\pi_m)^r+\frac{\pi_m}{\kappa}\sum_{j=1}^{r-1}j(1-\pi_m)^j+\frac{r}{\kappa}(1-\pi_m)^r\right]\\
&= Nc_1\left[1+\frac{(1-\pi_m)[1-(1-\pi_m)^{r-1}(1+\pi_m(r-1))]}{\pi_m\kappa}+\frac{r}{\kappa}(1-\pi_m)^r\right]\\
&= Nc_1\left[1+\frac{(1-\pi_m)[1-(1-\pi_m)^r]}{\pi_m\kappa}\right]\\
&= Nc_1\left[1+\frac{\pi_M\left[[1-(1-\pi_M)^{1/r}]^{-1}-1\right]}{\kappa}\right]
\end{aligned}$$

where

$$\pi_M = 1-(1-\pi_m)^r$$

is the proportion of subjects lost at the end of follow-up.

In addition, as expected, if there is no possibility of missing data, then

$$\begin{aligned}
\lim_{\pi_m\to 0}\mathbb{E}(\text{Cost}) &= Nc_1\lim_{\pi_m\to 0}\left[1+\frac{(1-\pi_m)[1-(1-\pi_m)^r]}{\pi_m\kappa}\right]\\
&= Nc_1\lim_{\pi_m\to 0}\left[1+\frac{(1-\pi_m)[1-(1-\pi_m r)]}{\pi_m\kappa}\right]\\
&= Nc_1\lim_{\pi_m\to 0}\left[1+\frac{(1-\pi_m)r}{\kappa}\right] = Nc_1\left(1+\frac{r}{\kappa}\right).
\end{aligned}$$

C Obtaining r_{opt} under the CMD response pattern in the basic scenario

Let r_{opt} be the value of r that minimizes the function

$$f(r,\kappa,\rho,\rho_e) = \frac{(\kappa+r)(1+\rho\, r)}{(r+1)(1-\rho+\rho(1-\rho_e)r)}, \tag{C.1}$$

with $r \in \mathbb{N}$, $\kappa \geqslant 1$, $\rho \in (0,1)$ and $\rho_e \in (0,1]$.

C.1 Time-invariant exposure ($\rho_e = 1$)

If $\rho_e = 1$, (C.1) reduces to

$$f(r,\kappa,\rho,\rho_e) = \frac{(\kappa+r)(1+\rho\, r)}{(1-\rho)(r+1)}$$

and we can equivalently minimize the function

$$g(r,\kappa,\rho,\rho_e) = \frac{(\kappa+r)(1+\rho\, r)}{r+1}. \tag{C.2}$$

If $\kappa = 1$, (C.2) reduces to

$$g(r,\kappa,\rho,\rho_e) = 1+\rho\, r$$

which is monotone increasing and then $r_{\text{opt}} = 0$.

If $\kappa > 1$, the minimum of (C.2) is allocated at $r_0 = -1 + \sqrt{\frac{1-\rho}{\rho}(\kappa-1)}$ which is greater than 0 if $\kappa > \frac{1}{1-\rho}$.

Then for time-invariant exposure:

$$r_{\text{opt}} = \begin{cases} 0 & , \quad \kappa \leqslant \frac{1}{1-\rho} \\ -1+\sqrt{\frac{1-\rho}{\rho}(\kappa-1)} & , \quad \kappa > \frac{1}{1-\rho} \end{cases} .$$

In general, r_0 is irrational. In that case, r_{opt} will be $[r_0]$ or $[r_0]+1$ ($[\cdot]$ is the integer part function), depending on which provides a lower value of $g(r,\kappa,\rho,\rho_e)$. Our R package take this correction into account.

C.2 Time-varying exposure ($0 < \rho_e < 1$)

The form of (C.1) changes depending on the values of κ, ρ and ρ_e. Hence, we consider different scenarios.

C.2.1 Case $\kappa = 1$

- If $\kappa = 1$ and $\rho_e = \rho$, (C.1) reduces to the constant $f(r,\kappa,\rho,\rho_e) = \frac{1}{1-\rho}$, independent of r and hence, for simplicity of the study design, we fix $r_{\text{opt}} = 0$.

- If $\kappa = 1$ and $\rho_e \neq \rho$, (C.1) reduces to

$$f(r, \kappa, \rho, \rho_e) = \frac{1 + \rho\, r}{1 - \rho + \rho(1 - \rho_e) r}$$

  and

$$\frac{\partial f(r, \kappa, \rho, \rho_e)}{\partial r} = \frac{\rho(\rho_e - \rho)}{[1 - \rho + \rho(1 - \rho_e) r]^2}$$

  so $f(r, \kappa, \rho, \rho_e)$ is monotone increasing if $\rho_e > \rho$ and monotone decreasing if $\rho_e < \rho$, for all positive value of r. Thus, $r_{\text{opt}} = 0$ if $\rho_e > \rho$ and $r_{\text{opt}} = +\infty$ if $\rho_e < \rho$.

C.2.2 Case $\boldsymbol{\kappa > 1}$

- If $\rho_e = \rho$, (C.1) reduces to

$$f(r, \kappa, \rho, \rho_e) = \frac{r + \kappa}{(1 - \rho)(r + 1)}$$

  and

$$\frac{\partial f(r, \kappa, \rho, \rho_e)}{\partial r} = -\frac{\kappa - 1}{(1 - \rho)(r + 1)^2} < 0 \quad \forall r \neq -1$$

  so $f(r, \kappa, \rho, \rho_e)$ is monotone decreasing and $r_{\text{opt}} = +\infty$.

- If $\rho_e \neq \rho$, there are still several different structures of $f(r, \kappa, \rho, \rho_e)$, related to the parameter

$$\kappa^\star := \frac{1 - \rho}{\rho(1 - \rho_e)} > 0.$$

  - If $\kappa = \kappa^\star > 1$ and $\rho_e \neq \rho$, (C.1) reduces to

$$f(r; \rho, \kappa, \kappa^\star) = \frac{\rho\, r + 1}{\rho(1 - \rho_e)(r + 1)}$$

    and

$$\frac{\partial f(r; \rho, \kappa, \kappa^\star)}{\partial r} = -\frac{1 - \rho}{\rho(1 - \rho_e)(r + 1)^2} < 0 \quad \forall r \neq -1\,,$$

    so $f(r)$ is monotone decreasing and $r_{\text{opt}} = +\infty$.
  - If $\kappa > 1$, $\kappa \neq \kappa^\star$ and $\rho_e \neq \rho$ (the general, most common case), (C.1) can be writen as

$$f(r; \rho, \kappa, \kappa^\star) = \frac{\kappa^\star}{1 - \rho} \cdot \frac{(r + \kappa)(1 + \rho\, r)}{(r + 1)(r + \kappa^\star)} \tag{C.3}$$

    and

$$\frac{\partial f(r; \rho, \kappa, \kappa^\star)}{\partial r} = \frac{\kappa^\star}{1 - \rho} \cdot \frac{[\rho(\kappa^\star - \kappa) - (1 - \rho)] r^2 + 2(\rho\kappa^\star - \kappa) r + [(\kappa^\star - \kappa) - \kappa\kappa^\star(1 - \rho)]}{(r + 1)^2 (r + \kappa^\star)^2}, \tag{C.4}$$

    so solving

$$\frac{\partial f(r; \rho, \kappa, \kappa^\star)}{\partial r} = 0$$

is equivalent to solving the equation

$$[\rho(\kappa^\star - \kappa) - (1-\rho)]r^2 + 2(\rho\kappa^\star - \kappa)r + [(\kappa^\star - \kappa) - \kappa\kappa^\star(1-\rho)] = 0. \tag{C.5}$$

If $\rho(\kappa^\star - \kappa) - (1-\rho)$ is equal to zero (equivalently, $\kappa = \frac{\rho_e(1-\rho)}{\rho(1-\rho_e)}$), then

$$\frac{\partial f(r;\rho,\kappa,\kappa^\star)}{\partial r} = -(\kappa - 1)\frac{2r + \kappa^\star(\kappa^\star + 1)}{(r+1)^2(r+\kappa^\star)^2} < 0 \quad \forall r \notin \{-\kappa^\star, -1\}$$

and $r_{\text{opt}} = +\infty$.

We consider now $\rho(\kappa^\star - \kappa) - (1-\rho) \neq 0$. The discriminant of (C.5) can be writen as

$$\Delta = 4(1-\rho)(\kappa - 1)(\kappa - \kappa^\star)(1 - \rho\kappa^\star)$$

and the sign of Δ is equal to the sign of $(\kappa - \kappa^\star)(1 - \rho\kappa^\star)$. Since $\kappa \neq \kappa^\star$ and $\rho\kappa^\star \neq 1$ (consequence of $\rho_e \neq \rho$), Δ can't be zero. Now, we analyze all the sign combinations of $(\kappa - \kappa^\star)$ and $(1 - \rho\kappa^\star)$:

* $1 - \rho\kappa^\star < 0$ (equivalently, $\rho_e > \rho$) and $\kappa > \kappa^\star$:
  In this case, $\Delta < 0$ and hence $f(r)$ is monotone $\forall r \notin \{-\kappa^\star, -1\}$. Then, since from (C.4),
  $$\frac{\partial f(r;\rho,\kappa,\kappa^\star)}{\partial r}\bigg|_{r=0} = -\frac{(\kappa - \kappa^\star) + \kappa\kappa^\star(1-\rho)}{(1-\rho)\kappa^\star} < 0,$$
  $f(r)$ is monotone decreasing $\forall r \notin \{-\kappa^\star, -1\}$ and $r_{\text{opt}} = +\infty$.
* $1 - \rho\kappa^\star > 0$ (equivalently, $\rho_e < \rho$) and $\kappa < \kappa^\star$:
  In this case, $\Delta < 0$ and hence $f(r)$ is monotone $\forall r \notin \{-\kappa^\star, -1\}$. Then, since from (C.4),
  $$\lim_{r\to+\infty} \frac{\partial f(r;\rho,\kappa,\kappa^\star)}{\partial r} = \lim_{r\to+\infty} \frac{\kappa^\star[\rho(\kappa^\star - \kappa) - (1-\rho)]}{(1-\rho)r^2},$$
  $f(r)$ is monotone decreasing $\forall r \notin \{-\kappa^\star, -1\}$, and then $r_{\text{opt}} = +\infty$, because $\rho(\kappa^\star - \kappa) - (1-\rho) < 0$ as shown by the following:
  $$1 - \rho\kappa^\star > 0 \Rightarrow 1 - \rho\kappa^\star > \rho(1-\kappa) \Rightarrow \rho(\kappa^\star - \kappa) - (1-\rho) < 0.$$
* $1 - \rho\kappa^\star > 0$ (equivalently, $\rho_e < \rho$) and $\kappa > \kappa^\star$:
  In this case, $\Delta > 0$ and the two solutions of (C.5) are
  $$r_\pm = \frac{\pm\frac{\sqrt{\Delta}}{2} - (\kappa - \rho\kappa^\star)}{(1-\rho) + \rho(\kappa - \kappa^\star)}$$
  whose denominator is positive and then $r_+ > r_-$. Further, since $\kappa > \kappa^\star$, then $r_- < 0$. From (C.4),
  $$\frac{\partial^2 f(r;\rho,\kappa,\kappa^\star)}{\partial r^2}\bigg|_{r=r_\pm} = \mp\frac{\kappa^\star\sqrt{\Delta}}{(1-\rho)(r_\pm + 1)^2(r_\pm + \kappa^\star)^2}, \tag{C.6}$$
  and there is a local minimum at $r_- < 0$ and a local maximum at r_+ so $f(r)$ is monotone decreasing for all $r > r_+$. We can show that $r_+ < 0$ and then $f(r)$ is

decreasing in $[0,+\infty)$ so $r_{\rm opt}=+\infty$. To show that $r_+<0$, we only need to prove that the numerator of r_+ is negative:

$$\begin{aligned}
\frac{\sqrt{\Delta}}{2}-(\kappa-\rho\kappa^\star) &= \sqrt{(1-\rho)(\kappa-1)(\kappa-\kappa^\star)(1-\rho\kappa^\star)}-(\kappa-\rho\kappa^\star)\\
&< \sqrt{(1-\rho)(\kappa-1)(\kappa-\rho\kappa^\star)(1-\rho\kappa^\star)}-(\kappa-\rho\kappa^\star)\\
&< \sqrt{(1-\rho)(\kappa-\rho\kappa^\star)(\kappa-\rho\kappa^\star)(1-\rho\kappa^\star)}-(\kappa-\rho\kappa^\star)\\
&= (\kappa-\rho\kappa^\star)\left[\sqrt{(1-\rho)(1-\rho\kappa^\star)}-1\right]\\
&< (\kappa-\rho\kappa^\star)\left[1-1\right]=0.
\end{aligned}$$

* $1-\rho\kappa^\star<0$ (equivalently, $\rho_e>\rho$) and $\kappa<\kappa^\star$:
In this case, $\Delta>0$ and the two solutions of (C.5) can be writen as

$$r_\pm=\frac{\pm\frac{\sqrt{\Delta}}{2}+(\rho\kappa^\star-\kappa)}{(1-\rho)-\rho(\kappa^\star-\kappa)} \tag{C.7}$$

and, by (C.6), there is a local minimum at $r_-<0$ and a local maximum at r_+ so $f(r)$ is monotone decreasing for all $r>r_+$. But now the denominator of (C.7) can be both positive and negative, so we will consider these two possibilities.
If the denominator of (C.7) is positive, then $r_+>r_-$ and $\kappa^\star<\kappa+\frac{1-\rho}{\rho}$ so

$$\begin{aligned}
\frac{\sqrt{\Delta}}{2}+(\rho\kappa^\star-\kappa) &= \sqrt{(1-\rho)(\kappa-1)(\kappa^\star-\kappa)(\rho\kappa^\star-1)}+\rho\kappa^\star-\kappa\\
&< \sqrt{(1-\rho)(\kappa-1)\left(\frac{1-\rho}{\rho}\right)(\rho\kappa-\rho)}+\rho\kappa+1-\rho-\kappa\\
&= \sqrt{(1-\rho)^2(\kappa-1)^2}+(1-\rho)(1-k)\\
&= (1-\rho)^2(\kappa-1)-(1-\rho)(k-1)=0
\end{aligned}$$

and then $r_+<0$ so $f(r)$ is monotone decreasing for all positive r and $r_{\rm opt}=+\infty$.
If the denominator of (C.7) is negative, then $r_+<r_-$ and $\kappa^\star>\kappa+\frac{1-\rho}{\rho}$. In this case, the sign of r_- is not constant. For example, if we have $\rho=0.55$ and $\rho_e=0.80$, if $\kappa=1.3$, the local mininum is at $r_-\approx-0.24$ so $r_{\rm opt}=0$ while if $\kappa=2.6$, the local minimum is at $r_-\approx4.08$ so $r_{\rm opt}=4$ or $r_{\rm opt}=5$. Then

$$r_{\rm opt}=\max\left(0,\frac{\sqrt{(1-\rho)(\kappa-1)(\kappa^\star-\kappa)(\rho\kappa^\star-1)}-\rho\kappa^\star+\kappa}{\rho(\kappa^\star-\kappa)-(1-\rho)}\right).$$

In addition, it can be shown that $\frac{\sqrt{(1-\rho)(\kappa-1)(\kappa^\star-\kappa)(\rho\kappa^\star-1)}-\rho\kappa^\star+\kappa}{\rho(\kappa^\star-\kappa)-(1-\rho)}<0$ if $\kappa\leqslant\frac{\rho\kappa^\star}{1+(1-\rho)\rho\kappa^\star}$ and thus,

$$r_{\rm opt}=\begin{cases} 0 & ,\ \text{if } \kappa\leqslant\frac{\rho\kappa^\star}{1+(1-\rho)\rho\kappa^\star}\\ \frac{\sqrt{(1-\rho)(\kappa-1)(\kappa^\star-\kappa)(\rho\kappa^\star-1)}-\rho\kappa^\star+\kappa}{\rho(\kappa^\star-\kappa)-(1-\rho)} & ,\ \text{otherwise}\end{cases}$$

C.3 Summary results for the CMD response pattern

The results are summarized in Table C.1.

Table C.1: Optimal total number of measurements under the CMD response pattern in the basic scenario.

Time-varying exposure ($\rho_e < 1$)		
$\boldsymbol{\rho_e}$	$\boldsymbol{\kappa}$	**Optimal total number of measurements**
$\rho_e \geq \rho$	1	1
$\rho_e > \rho$	$\kappa \leq \kappa_0$	1
	$\kappa \in (\kappa_0, \kappa_c)$	$1 + r_0$
other combinations of $(\kappa,\ \rho,\ \rho_e)$		$+\infty$

where:

$\kappa_0 := \frac{1-\rho}{1-\rho_e+(1-\rho)^2}$, $\kappa_c := \frac{\rho_e(1-\rho)}{\rho(1-\rho_e)}$

$r_0 := \frac{\kappa-1-\rho(\kappa_c-1)+\sqrt{(1-\rho)(\kappa_c-1)(\kappa-1)[1-\rho+\rho(\kappa_c-\kappa)]}}{\rho(\kappa_c-\kappa)}$

Time-invariant exposure ($\rho_e = 1$)	
$\boldsymbol{\kappa}$	**Optimal total number of measurements**
$\kappa \leqslant \frac{1}{1-\rho}$	1
Otherwise	$\sqrt{\frac{1-\rho}{\rho}(\kappa-1)}$

D Obtaining r_{opt} under the LDD response pattern in the basic scenario

Let r_{opt} be the value of r that minimizes the function

$$f(r,\kappa,\rho,\rho_e) = \frac{r(r+\kappa)(\rho r+1)}{(r+1)[\rho\rho_e r^2 + [2\rho+\rho_e+3(1-\rho)\rho_e(1-\rho_e)]r + 2[1+(2-\rho)(1-\rho_e)]]}, \tag{D.1}$$

with $r \in \mathbb{N}^+$, $\kappa \geqslant 1$, $\rho \in (0,1)$ and $\rho_e \in (0,1]$.

D.1 Time-invariant exposure ($\rho_e = 1$)

If $\rho_e = 1$, (D.1) reduces to

$$f(r,\kappa) = \frac{r(r+\kappa)}{(r+1)(r+2)} \tag{D.2}$$

which does not depend on ρ.

- If $\kappa = 1$ or $\kappa = 2$, (D.2) reduces to

  $$f(r,\kappa) = \frac{r}{r+3-\kappa}$$

  and

  $$\frac{\partial f(r,\kappa)}{\partial r} = \frac{3-\kappa}{(r+3-\kappa)^2} > 0 \quad \forall r \notin \{-2,-1\}$$

  so $f(r,\kappa)$ is monotone increasing and $r_{\text{opt}} = 1$.

- If $\kappa \notin \{1,2\}$,

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

  so solving

  $$\frac{\partial f(r,\kappa)}{\partial r} = 0$$

  is equivalent to solving the equation

  $$(3-\kappa)r^2+4r+2\kappa = 0. \tag{D.4}$$

  - If $\kappa = 3$,

    $$\frac{\partial f(r,\kappa)}{\partial r} = \frac{4r+2\kappa}{(r+1)^2(r+2)^2} > 0 \quad \forall r \geqslant 0$$

    so $f(r,\kappa)$ is monotone increasing and $r_{\text{opt}} = 1$.

  - If $\kappa \neq 3$, the solution of (D.4) is

    $$r_{\pm} = \frac{-2 \pm \sqrt{2(\kappa-1)(\kappa-2)}}{3-\kappa}. \tag{D.5}$$

If $\kappa < 2$ the discriminant of (D.5) is negative and hence $f(r,\kappa)$ is monotone. And, since from (D.3) $\frac{\partial f(r,\kappa)}{\partial r}$ is positive at $r = 0$, $f(r,\kappa)$ is monotone increasing and $r_{\mathrm{opt}} = 1$.
If $\kappa > 2$, both values of (D.5) are real.
If $\kappa \in (2,3)$, from (D.5) we have

$$r_- < r_+ = \frac{-2 + \sqrt{2(\kappa-1)(\kappa-2)}}{3-\kappa} < \frac{-2+\sqrt{2\cdot 2\cdot 1}}{3-\kappa} = 0\,,$$

so $f(r,\kappa)$ is monotone increasing for non negative r and hence $r_{\mathrm{opt}} = 1$.
If $\kappa > 3$, $r_+ < 0$ and $r_- > 0$ so we only need to analize r_-. By differentiating (D.3) and given that $r_\pm$ is solution of (D.4), we have

$$\frac{\partial^2 f(r,\kappa)}{\partial r^2}\bigg|_{r=r_\pm} = \pm\frac{2\sqrt{2(\kappa-1)(\kappa-2)}}{(r_\pm+1)^2(r_\pm+2)^2} \tag{D.6}$$

which is negative for $r = r_-$, so $f(r,\kappa)$ has a local maximum at $r = r_-$ and $f(r,\kappa)$ is increasing for $0 \leqslant r < r_-$ and decreasing for $r > r_-$. On the other hand, it can easily be shown that r_- is decreasing in κ for $\kappa > 3$ and that $\lim_{\kappa\to+\infty} r_- = \sqrt{2}$, so $r_- > 1$ for all $\kappa > 3$. Hence, $r_{\mathrm{opt}} = 1$ if $f(r=1,\kappa) \leqslant \lim_{r\to+\infty} f(r,\kappa)$ (equivalently, from (D.2), to $\kappa \leqslant 5$) and $r_{\mathrm{opt}} = +\infty$ otherwise.

In summary, for a time-invariant exposure:

$$r_{\mathrm{opt}} = \begin{cases} 1 & , \quad \kappa \leqslant 5 \\ +\infty & , \quad \kappa > 5 \end{cases} .$$

D.2 Time-varying exposure ($\rho_e < 1$)

D.2.1 Case $\kappa = 1$

In this case, (D.1) reduces to:

$$f(r;\kappa,\rho,\rho_e) = \frac{r(\rho r+1)}{\rho\rho_e r^2 + [2\rho+\rho_e+3(1-\rho)\rho_e(1-\rho_e)]r + 2[1+(2-\rho)(1-\rho_e)]}\,. \tag{D.7}$$

Note that since all three coefficients in the polynomial of degree 2 in r in the denominator of $f(r)$ are positive, $f(r;\kappa,\rho,\rho_e)$ is continuous for all positive r.

Also,

$$\frac{\partial f(r;\kappa,\rho,\rho_e)}{\partial r} = \frac{\rho[2\rho+3(1-\rho)\rho_e(1-\rho_e)]r^2 + 4\rho[1+(2-\rho)(1-\rho_e)]r + 2[1+(2-\rho)(1-\rho_e)]}{[\rho\rho_e r^2 + [2\rho+\rho_e+3(1-\rho)\rho_e(1-\rho_e)]r + 2[1+(2-\rho)(1-\rho_e)]]^2}. \tag{D.8}$$

Note now that all six coefficients in this rational function are positive. Hence, $\frac{\partial f(r;\kappa,\rho,\rho_e)}{\partial r}$ is positive for all positive r and thus $f(r;\kappa,\rho,\rho_e)$ is monotone increasing for all positive r, and $r_{\mathrm{opt}} = 1$.

D.2.2 Case $\kappa > 1$

We consider two cases: $\rho_e = \frac{2\rho(2-\rho)}{3(1-\rho)}$ and $\rho_e \neq \frac{2\rho(2-\rho)}{3(1-\rho)}$.

- Case $\rho_e = \frac{2\rho(2-\rho)}{3(1-\rho)}$ (which implies $\rho < \frac{1}{2}$):

  In this case, ignoring constant positive factors, (D.1) reduces to

  $$f(r;\kappa,\rho,\rho_e) = \frac{r(r+\kappa)}{(r+1)(r+\alpha)}, \qquad \alpha := \frac{9-20\rho+11\rho^2-2\rho^3}{\rho(2-\rho)} \in (+2,+\infty) \tag{D.9}$$

  If $\kappa=\alpha$, $f(r;\kappa,\rho,\rho_e) = \frac{r}{r+1}$ which is monotone increasing for all positive r and $r_{\text{opt}} = 1$.

  We consider now $\kappa \neq \alpha$. Then

  $$\frac{\partial f(r;\kappa,\rho,\rho_e)}{\partial r} = \frac{(\alpha+1-\kappa)r^2+2\alpha r+\alpha\kappa}{(r+1)^2(r+\alpha)^2}\,. \tag{D.10}$$

  If $\kappa \leqslant \alpha+1$, then $f(r;\kappa,\rho,\rho_e)$ is monotone increasing for all positive r, and $r_{\text{opt}} = 1$.

  If $\kappa > \alpha+1$, the roots of (D.10) are

  $$r_\pm = \frac{\alpha \pm \sqrt{\Delta}}{\kappa-(\alpha+1)}, \quad \text{where} \quad \Delta = \alpha(\kappa-1)(\kappa-\alpha) > \alpha^2 \tag{D.11}$$

  which implies $r_- < 0$ and we only need to analyze r_+ where there is a local maximum since

  $$\frac{\partial^2 f(r,\kappa,\rho,\rho_e)}{\partial r^2}\bigg|_{r=r_+} = \frac{-2\sqrt{\Delta}}{(r_+ +1)^2(r_+ +\alpha)^2} < 0.$$

  It can be shown that $r_+ > 1$ and then $f(r;\kappa,\rho,\rho_e)$ is increasing for $r \in [1,r_+)$ and decreasing for $r \in (r_+,+\infty)$. Thus, $r_{\text{opt}} = 1$ if $f(r=1;\kappa,\rho,\rho_e) \leqslant \lim_{r\to+\infty} f(r;\kappa,\rho,\rho_e)$ (equivalently, if $\kappa \leqslant 2\alpha+1$) and $r_{\text{opt}} = +\infty$ otherwise.

  Summarizing,

  $$\rho_e = \frac{2\rho(2-\rho)}{3(1-\rho)} \quad \Rightarrow \quad r_{\text{opt}} = \begin{cases} 1 & , \quad \kappa \leqslant \kappa^* \\ +\infty & , \quad \kappa > \kappa^* \end{cases} \quad , \qquad \kappa^* := \frac{18-38\rho+21\rho^2-4\rho^3}{\rho(2-\rho)}$$

- Case $\rho_e \neq \frac{2\rho(2-\rho)}{3(1-\rho)}$ (the most general and common case):

  In this case, the analytical optimization of (D.1) is practically impossible because it requires solving a polynomial equation of degree 4 with complex expressions for their coefficients, which depend on ρ, ρ_e and κ. Over a fine grid of values of the parameters $\rho \in (0,1)$, $\rho_e \in (0,1)$ and $\kappa \in (1, 10000)$, we found that $r_{\text{opt}} = 1$ or $r_{\text{opt}} = +\infty$ always. Assuming that the value of r_{opt} can be only 1 or $+\infty$, and since from (D.1), $f(r=1) = \frac{(\kappa+1)(1+\rho)}{6[2-(1-\rho)\rho_e^2]}$ and $\lim_{r\to+\infty} f(r) = \frac{1}{\rho_e}$, we have

  $$r_{\text{opt}} = \begin{cases} 1 & , \quad \kappa \leqslant \kappa^* \\ +\infty & , \quad \kappa > \kappa^* \end{cases} \quad , \qquad \kappa^\star := 5 + \frac{6(1-\rho_e)[2+(1-\rho)\rho_e]}{(1+\rho)\rho_e} \tag{D.12}$$

which is coherent with results in all the others cases analyzed in this section.

In addittion, since

$$\frac{\partial \kappa^*}{\partial \rho} = -\frac{12(1-\rho_e^2)}{(1+\rho)\rho_e} < 0 \quad , \qquad \frac{\partial \kappa^*}{\partial \rho_e} = -\frac{6[2+(1-\rho_e)\rho^2]}{(1+\rho)\rho_e^2} < 0\,,$$

κ^* decreases with both ρ and ρ_e.

D.3 Summary results for the LDD response pattern

The results are summarized in Table D.1. In general, we should take only one measurement if κ is not greater than the threshold $\frac{6[2-(1-\rho)\rho_e^2]}{(1+\rho)\rho_e} - 1$, which decreases with both ρ and ρ_e.

Table D.1: Optimal total number of measurements under the LDD response pattern in the basic scenario.

Time-varying exposure ($\rho_e < 1$)	
$\boldsymbol{\kappa}$	**Optimal total number of measurements**
$\kappa \leqslant \frac{6[2-(1-\rho)\rho_e^2]}{(1+\rho)\rho_e} - 1$	2
Otherwise	$+\infty$

Time-invariant exposure ($\rho_e = 1$)	
$\boldsymbol{\kappa}$	**Optimal total number of measurements**
$\kappa \leqslant 5$	2
Otherwise	$+\infty$

E Simulation study for the effect of dropout

In this simulation study, we assumed that the probability of dropout of participant i at measurement j was π_1 if the response at the measurement $j-1$ was lower than its third quartile or π_2 otherwise. We explored 9 scenarios, obtained from all combinations of the overall dropout fraction (0.1, 0.3 and 0.6) and the ratio π_1/π_2 (0.5, 0.8 and 1, corresponding this last value to missing completely at random, MCAR).

E.1 Parameterization

E.1.1 Parameters

- $\pi_m = P(Y_j = \text{missing})$
- $\pi_M = P(\text{loosing } i\text{-th participant})$
- $\pi_1 = P(Y_j = \text{missing}|Y_{j-1} < Y_{j-1}^{(p)})$
- $\pi_2 = P(Y_j = \text{missing}|Y_{j-1} \geqslant Y_{j-1}^{(p)})$
- $Y_{j-1}^{(p)}$ is the p-quantile of Y_{j-1}, i.e., $P(Y_{j-1} \leqslant Y_{j-1}^{(p)}) = p$
- $\lambda := \pi_1/\pi_2$.
- Input parameters: π_M and λ
- Simulation parameters: π_1 and π_2

E.1.2 Obtaining π_1 and π_2 as a function of π_M and λ

Assuming a monotone missing pattern,

$$\pi_M = 1 - (1 - \pi_m)^r.$$

Also,

$$\begin{aligned}
\pi_m &= P(Y_j = \text{missing}) \\
&= P(Y_j = \text{missing}|Y_{j-1} < Y_{j-1}^{(p)})\, P(Y_{j-1} < Y_{j-1}^{(p)}) \\
&+ P(Y_j = \text{missing}|Y_{j-1} \geqslant Y_{j-1}^{(p)})\, P(Y_{j-1} \geqslant Y_{j-1}^{(p)}) \\
&= \pi_1\, p + \pi_2(1-p) = \pi_2[1 - (1-\lambda)p].
\end{aligned}$$

Thus,

$$\begin{cases} \pi_2(\pi_M, \lambda)_{(r,p)} = \dfrac{1 - (1-\pi_M)^{1/r}}{1 - (1-\lambda)p} \\ \pi_1(\pi_M, \lambda)_{(r,p)} = \lambda\, \pi_2 \end{cases} \tag{E.1}$$

From (E.1), must be

$$\lambda \geqslant 1 - \frac{(1-\pi_M)^{1/r}}{p}.$$

E.2 Scenarios

We set some parameters to fixed values (most of them from the illustrative example in the paper): $N = 43$, $r = 14$, $\theta = 0$, $\rho = 0.3$, $\rho_e = 0.13$, $p_{e0} = p_{er} = 0.37$, $\sigma^2 = 0.43$, $p = 0.75$, $\beta_0 = 1$, $\beta_{\text{time}} = -0.5$, $\beta_{\text{E}_0} = -0.5$ (for LDD model), $\alpha = 0.05$. In each scenario, the value of the β of interest was set to that value which provides an expected power according to our formulas of 0.8. The number of simulations per scenario was 500.

E.3 Results

	Empirical power ($\mathbf{CI_{95\%}}$)		
$\boldsymbol{\pi_M}$	$\boldsymbol{\lambda = 0.5}$	$\boldsymbol{\lambda = 0.8}$	$\boldsymbol{\lambda = 1}$ **(MCAR)**
0.1	0.79 (0.75, 0.82)	0.81 (0.77, 0.84)	0.80 (0.77, 0.84)
0.3	0.82 (0.79, 0.86)	0.79 (0.76, 0.83)	0.80 (0.77, 0.83)
0.6	0.79 (0.76, 0.83)	0.80 (0.77, 0.84)	0.77 (0.73, 0.81)

Table E.1: CMD model: Empirical power when expected power was 0.8.

	Empirical power ($\mathbf{CI_{95\%}}$)		
$\boldsymbol{\pi_M}$	$\boldsymbol{\lambda = 0.5}$	$\boldsymbol{\lambda = 0.8}$	$\boldsymbol{\lambda = 1}$ **(MCAR)**
0.1	0.82 (0.79, 0.86)	0.76 (0.72, 0.80)	0.82 (0.78, 0.85)
0.3	0.77 (0.73, 0.80)	0.80 (0.77, 0.84)	0.78 (0.74, 0.81)
0.6	0.77 (0.73, 0.80)	0.79 (0.76, 0.83)	0.80 (0.77, 0.84)

Table E.2: LDD model: Empirical power when expected power was 0.8.

Results of the simulation study are shown in Table E.1 (CMD model) and Table E.2 (LDD model). Any scenario showed significant differences in terms of power when comparing the empirical values with the expected one.

F Effect of a time-varying exposure prevalence on the results, in the basic scenarios

Table F.1: Effect of time-varying exposure prevalence on r_{opt} for the CMD response pattern and CS(ρ) covariance structure of the response, no missingness and CS(ρ_e) exposure covariance structure. Exploration was performed for $r \leqslant 30$ and for all combinations of the parameters $\rho_e = 0.3, 0.8$, $\rho = 0.2, 0.8$, $\kappa = 1, 2, 5, 8$, $\bar{p}_e = 0.1, 0.3, 0.5$ and $p_{er}/p_{e0} = 1/20, 1, 21$ (i.e., $\gamma = -0.95, 0, 20$, respectively). Numbers in the body of the table correspond to r_{opt}.

		$\rho_e = 0.3$			$\rho_e = 0.8$					
		$\rho = 0.2$		$\rho = 0.8$	$\rho = 0.2$				$\rho = 0.8$	
p_{er}/p_0	$\bar{p}_e$	$\kappa = 1$	$\kappa \geqslant 2$	$\kappa \geqslant 1$	$\kappa = 1$	$\kappa = 2$	$\kappa = 5$	$\kappa = 8$	$\kappa = 1$	$\kappa \geqslant 2$
	0.1	0	30	30	0	0	7	17	1	30
1/20	0.3	0	30	30	0	0	8	16	0	30
	0.5	30	30	30	4	5	10	19	6	30
1	$\star$	0	30	30	0	2	6	11	6	30
	0.1	2	30	30	1	2	7	17	1	30
21	0.3	8	30	30	2	3	8	16	2	30
	0.5	30	30	30	4	5	10	19	6	30

$\star$: r_{opt} does not depend on $\bar{p}_e$ if the prevalence of exposure is constant.

Table F.2: Effect of time-varying exposure prevalence on r_{opt} for the LDD response pattern and CS(ρ) response covariance structure, no missingness and CS(ρ_e) exposure covariance structure. Exploration was performed for $r \leqslant 30$ and for all combinations of the parameters $\rho_e = 0.3, 0.8$, $\rho = 0.2, 0.8$, $\kappa = 1, 2, 5, 8$, $\bar{p}_e = 0.1, 0.3, 0.5$ and $p_{er}/p_{e0} = 1/20, 1, 21$ (i.e., $\gamma = -0.95, 0, 20$, respectively). Numbers in the body of the table correspond to r_{opt}.

		$\rho_e = 0.3$						$\rho_e = 0.8$							
		$\rho = 0.2$			$\rho = 0.8$			$\rho = 0.2$				$\rho = 0.8$			
p_{er}/p_0	$\bar{p}_e$	$\kappa \leqslant 2$	$\kappa = 5$	$\kappa = 8$	$\kappa \leqslant 2$	$\kappa = 5$	$\kappa = 8$	$\kappa = 1$	$\kappa = 2$	$\kappa = 5$	$\kappa = 8$	$\kappa = 1$	$\kappa = 2$	$\kappa = 5$	$\kappa = 8$
	0.1	2	← – 3 – →		2	3	4	3	4	← – 30 – →		3	5	← – 30 – →	
1/20	0.3	← – 2 – →		3	2	← – 3 – →		2	3	← – 30 – →		3	4	← – 30 – →	
	0.5	← – – – 2 – – – →			← – – – 2 – – – →			← – 2 – →		26	30	2	3	← – 30 – →	
1	$\star$	← – – – – – – – – – – – – – 1 – – – – – – – – – – – – – →													30
	0.1	← – – – – – – – – – – – – – – – 1 – – – – – – – – – – – – – – – →													
21	0.3	← – – – – – – – – – – – – – – – 1 – – – – – – – – – – – – – – – →													
	0.5	← – – – – – – – – – – 2 – – – – – – – – – – →								← – 30 – →		2	3	← – 30 – →	

$\star$: r_{opt} does not depend on $\bar{p}_e$ if the prevalence of exposure is constant.

G Simulation study for the effect of confounding

G.1 Simulation study 1

A simulation study was performed in order to explore the impact of confounding. Scenarios explored corresponded to all combinations of the values of the parameters: $\rho_e = 0.2, 0.8$; $N = 50, 200, 500, 2000$; and $r = 1$ (only for CMD), 2 (only for LDD), $8, 20$. Each of these scenarios was explored for five confounding patterns and for a number of confounders from 0 to 6. Thus, a total of 840 scenarios were explored. The remaining parameters were fixed at the values: $\sigma^2 = 1$, $\theta = 0$, $\rho = 0.5$, $\alpha = 0.05$, $p_e = 0.3$ and $\pi_M = 0$. The β's of the model were fixed at 0.5 times the value of the β of interest. The value of the β of interest was obtained according to our formulas in order to achieve a power of 0.9

The simulation study was performed according to the following steps:

1. Up to 6 simulated confounders were considered, in the following order:

$$
\begin{aligned}
Z_{1j} &\sim \mathrm{N}(0,1), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{1j}, Z_{1k}) = \rho_1 \text{ for } k \neq j, \\
Z_{2j} &\sim \mathrm{Bernoulli}(p_2), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{2j}, Z_{2k}) = \rho_2 \text{ for } k \neq j, \\
Z_{3j} &\sim \mathrm{N}(0,1), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{3j}, Z_{3k}) = \rho_3 \text{ for } k \neq j, \\
Z_{4j} &\sim \mathrm{Bernoulli}(p_4), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{4j}, Z_{4k}) = \rho_4 \text{ for } k \neq j, \\
Z_{5j} &\sim \mathrm{N}(0,1), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{5j}, Z_{5k}) = \rho_5 \text{ for } k \neq j, \\
Z_{6j} &\sim \mathrm{Bernoulli}(p_6), \quad j = 0, \dots, r, \text{ with } \mathrm{corr}(Z_{6j}, Z_{6k}) = \rho_6 \text{ for } k \neq j,
\end{aligned}
$$

   and assuming independence across the Z_{mj}. p_m was set to 0.2, $\forall m = 2, 4, 6$, and ρ_m was set to 0.5 $\forall m = 1, 2, \dots, 6$.

2. Exposure E was simulated as a multivariate correlated Bernoulli random variable with parameters p_e and ρ_e.

3. Each of the binary confounders $Z_{mj}, \forall m = 2, 4, 6$, was simulated from the logistic model

$$\log\left(\frac{P(Z_m = 1)}{P(Z_m = 0)}\right) = \alpha_{0m} + E \, \log\phi_m,$$

   where ϕ_m is the odds ratio of E associated to Z_m, and α_{0m} was fixed from the constraint $\mathbb{E}(Z_m) = p_m$.

4. Each of the normal confounders $Z_{mj}, \forall m = 1, 3, 5$, was simulated from the normal distribution

$$Z_m \sim \mathrm{N}(\mu_m = E \log\phi_m, \sigma_m = 1),$$

   where ϕ_m is the odds ratio of E associated to an one unit increase of Z_m.

5. The value of the β of interest was obtained according to our formulas in order to achieve a power of 0.9.

6. Multivariate normal response was simulated with mean equal to the simulated linear predictor based on models (1) and (2), including the first q confounders, with q between 0 and 6. The response covariance structure was DEX(σ, θ, ρ).

   Five confounding scenarios were considered for the association between the confounders and both the exposure and the response. These scenarios are described in Table G.1.

Confounding scenario	**OR between Z_{mj} and X_j, ϕ_m**	**Effect of Z_{mj} on Y_j, β_{Z_m} †**
Constant strength		
1. Weak	1.5	0.5
2. Moderate	2.0	1.0
3. Strong	2.5	1.5
Diminish strength		
4. Moderate	$3.0 \to 1.5$‡	$2.0 \to 0.5$
5. Strong	$4.0 \to 1.5$‡	$3.0 \to 0.5$

†: In units of the β of interest.
‡: Linearly decreasing with the number of confounders.

Table G.1: Confounding scenarios.

7. Then, the model was fitted to the simulated data and the p-value for the significance of the β of interest (F-test) was stored.

8. Steps 1 to 7 were iterated 1000 times.

9. The empirical power was obtained as the fraction of times that p-value $\leqslant \alpha$. Point estimate and percentiles 2.5 and 97.5 were stored.

10. Steps 1 to 9 were performed for all combinations of $q = 0, \ldots, 6$, the five confounding scenarios and the explored values for ρ_e, N, and r.

11. A total of 840 scenarios were explored.

Results under the CMD response pattern are showed in Figures G.1 and G.1 in scenarios with $\rho_e = 0.2$ and $\rho_e = 0.8$, respectively. Under the LDD response pattern, results are shown in Figures G.3 to **??**.

G.2 Simulation study 2

In addition, we performed another simulation study in order to explore the potential effect of confounding on the optimal combination of N and r. We explored 8 scenarios, combining the values of ρ_e (0.2 and 0.8), the ratio of costs between the first measurement and the subsequent ones, κ (2 and 5 under the model (1); 2 and 10 under the model (2)) and the confounding strength of the confounders ("constant weak" and "constant strong"). For each scenario, the values of N and r were fixed at the optimal values according to our formulas when minimizing the cost of the study while achieving a power of 0.9. Then, the empirical optimal combination of N and r was fixed as that combination which provided the maximum empirical power without exceeding the minimum cost mentioned above. Empirical results were assessed by including 4 confounders and simulating 1000 datasets for each scenario. Results showed an impact of confounding on the study design only for higher values of ρ_e and κ. Thus, under the LDD model, discrepancies were observed only for $\rho_e = 0.8$ and $\kappa = 10$, with empirical $r_{\text{opt}} = 13$ and $N_{\text{opt}} = 80$ when expecting $r_{\text{opt}} = 20$ and $N_{\text{opt}} = 62$. Under the CMD model, similar results were obtained, which are detailed in Table G.2.

Table G.2: Impact of confounding on the optimal combination of number of repeated measurements and number of participants under the CMD response pattern.

$\boldsymbol{\rho_e}$	$\boldsymbol{\kappa}$	**Confounding strength**†	$(\boldsymbol{r_{\text{opt}}}, \boldsymbol{N_{\text{opt}}})$ **Expected**	**Empirical**
0.2	2	Weak	(18, 7)	(15, 8)
		Strong		(13, 9)
	5	Weak	(18, 7)	(18, 7)
		Strong		(18, 7)
0.8	2	Weak	(1, 126)	(2, 94)
		Strong		(1, 126)
	5	Weak	(20, 21)	(11, 32)
		Strong		(11, 32)

† See Table G.1.

H R package usage

In this section, we illustrate the usage of the R package `optimalAllocation` with some examples, including the reproduction of the results showed in the section 5 of the manuscript. The package can be downloaded at `http://www.creal.cat/xbasagana/software.html`.

H.1 Some examples

H.1.1 Study 1. Maximizing power

Suppose we are interested in maximizing the power of a longitudinal study assuming the CMD response pattern without exceeding a budget of 40 monetary units, where the monetary unit is the cost of the first measurement. The cost of the first measurement is $\kappa = 3$ times the cost of the subsequent ones. The response covariance structure is DEX($\sigma = 1, \rho = 0.7, \theta = 0.5$). The exposure intraclass correlation is $\rho_e = 0.2$. The expected proportion of dropout at the end of the study is $\pi_M = 0.2$. The exposure prevalence is assumed to increase linearly from $p_{e0} = 0.2$ at the first measurement to $p_{er} = 0.3$ at the last measurement. The effect size to be detected is $\beta = -0.3$ and the significance level is fixed at $\alpha = 0.05$. The maximum number of repeated measurements allowed is $r_{\max} = 20$. Thus, we can perform the study calculations and store the results in the object `study1`:

```
> library(optimalAllocation)
> study1 <- OA(target = "maxPower", pattern = "CMD", rMax = 20, theta = 0.5,
+              rho = 0.7, sigma = 1, rhoe = 0.2, pe0 = 0.2, per = 0.3, piM = 0.2,
+              kappa = 3, budget = 40, c1 = 1, beta = -0.3, alpha = 0.05)
>
> study1
  Results subject to r not greater than 20:
-----------------------------------------------------
Optimal total number of measurements (r+1): 20
Optimal number of participants (N)        : 6
Maximized power                           : 0.9670238
```

Thus, the optimal is to perform a longitudinal study with $N_{\text{opt}} = 6$ participants and taking $r_{\text{opt}} + 1 = 20$ measurements. The maximized power of such study is 0.97.

Further information can be obtained from the function `plot()`. For instance, Figure H.1 is the output of `plot(study1)` and shows that the optimal strategy is to take as many measurements as possible. Further results, including the estimated standard error of β, can be obtained with `summary(study1)`.

H.1.2 Study 2. Particular case: Cost of a cross-sectional study

Suppose we are interested in finding the cost of a cross-sectional study achieving a power of at least 0.9 to detect an effect size $\beta = -0.3$ with a significance level $\alpha = 0.05$. The cost of the unique measurement per participant is $c_1 = 25$ monetary units. The proportion of exposed is assumed to be 0.3 and the residual variance is estimated in $\sigma = 1$. Thus, the study calculations are:

```
> study2 <- OA(target = "minCost", pattern = "CMD", rMax = 0, sigma = 1,
+            pe0 = 0.3, reqPower = 0.9, c1 = 25, beta = -0.3, alpha = 0.05)
> study2
  Results subject to a cross-sectional design:
-----------------------------------------------------
Number of participants (N): 556
Cost                      : 13900
```

Thus, the required number of participants is $N = 556$ and the total cost is 13,900 monetary units.

H.2 Obtaining results in section 5

As another example of the package usage, we reproduce here results in the section 5 of the manuscript. In that example, $\rho_e = 0.13$ and the constant exposure prevalence was $p_{e0} = p_{er} = 0.37$ for vacuum cleaning, and $\rho_e = 0.60$ and $p_{e0} = p_{er} = 0.17$ for using air freshener sprays. The dropout fraction at the end of the study was $\pi_M = 0.28$. The residual variance and the response covariance damping parameter were taken from the study and set to $\sigma^2 = 0.43$ and $\theta = 0.12$, respectively. We used low (0.3) and high (0.7) values for ρ. The hypothesized effect was fixed at a difference of 10% in the expected mean value of the response between exposed and non exposed assuming the CMD response pattern (i.e., $\tilde{\beta} = -0.39$). The objective was to minimize the total cost of the study fixing a minimum required power of 0.9. The first measurement was assumed to be 2 times more expensive than each of the subsequent ones (i.e., $\kappa = 2$). We constrained the maximum number of repeated measurements to 20. All calculations were performed fixing a significance level $\alpha = 0.05$.

Then, results in Table 5 (section 5 of the manuscript) can be reproduced with the following code:

```
> # Creating scenarios:
>
> Table4 <- expand.grid(Exposure = c("Vacuuming", "Air freshener sprays"),
+                       rho = c(0.3, 0.7))
> Table4$pe0 <- 0.37
> Table4$pe0[Table4$Exposure == "Air freshener sprays"] <- 0.17
> Table4$per <- Table4$pe0
> Table4$rhoe <- 0.13
> Table4$rhoe[Table4$Exposure == "Air freshener sprays"] <- 0.60
> Table4$TimeVaryingExposure <- TRUE
> Table42 <- Table4
> Table42$TimeVaryingExposure <- FALSE
> Table42$rhoe <- 1
> Table4 <- rbind(Table4, Table42)
> Table4$r <- NA
> Table4$N <- NA
> Table4$cost <- NA
>
> # Sorting as in Table 4:
> ord <- order(Table4$Exposure, Table4$rho, 1 - Table4$TimeVaryingExposure)
```

```
> Table4 <- Table4[ord, ]
> rownames(Table4) <- NULL
>
> Table4
             Exposure rho  pe0  per rhoe TimeVaryingExposure  r  N cost
1           Vacuuming 0.3 0.37 0.37 0.13                TRUE NA NA   NA
2           Vacuuming 0.3 0.37 0.37 1.00               FALSE NA NA   NA
3           Vacuuming 0.7 0.37 0.37 0.13                TRUE NA NA   NA
4           Vacuuming 0.7 0.37 0.37 1.00               FALSE NA NA   NA
5 Air freshener sprays 0.3 0.17 0.17 0.60                TRUE NA NA   NA
6 Air freshener sprays 0.3 0.17 0.17 1.00               FALSE NA NA   NA
7 Air freshener sprays 0.7 0.17 0.17 0.60                TRUE NA NA   NA
8 Air freshener sprays 0.7 0.17 0.17 1.00               FALSE NA NA   NA
>
> # Optimal allocation calculations
> # for all scenarios:
>
> studies <- list()
>
> for (i in 1:nrow(Table4))
+  {
+   studies[[i]] <- OA(target = "minCost", pattern = "CMD", rMax = 20,
+                      theta = 0.12, rho = Table4$rho[i], sigma = sqrt(0.43),
+                      rhoe = Table4$rhoe[i], pe0 = Table4$pe0[i],
+                      per = Table4$per[i], piM = 0.28, kappa = 2,
+                      reqPower = 0.9, c1 = 1, beta = -0.39, alpha = 0.05)
+   Table4$r[i] <- studies[[i]]$ropt
+   Table4$N[i] <- studies[[i]]$Nopt
+   Table4$cost[i] <- round(studies[[i]]$minCost, 1)
+  }
>
> # Results:
>
> Table4[, -c(3:5)]
             Exposure rho TimeVaryingExposure  r   N  cost
1           Vacuuming 0.3                TRUE 18   6  51.6
2           Vacuuming 0.3               FALSE  1  92 125.1
3           Vacuuming 0.7                TRUE 15   3  22.0
4           Vacuuming 0.7               FALSE  0 128 128.0
5 Air freshener sprays 0.3                TRUE 20  17 160.7
6 Air freshener sprays 0.3               FALSE  1 152 206.7
7 Air freshener sprays 0.7                TRUE 19   8  72.2
8 Air freshener sprays 0.7               FALSE  0 211 211.0
```

In order to explore the effect of departures of the value of r from the value of r_{opt}, we can create the Figure H.2 using the following code:

```
> par(las=1, mfrow=c(2, 2))
> for (i in c(1,3,5,7))
>  {
>   plot(studies[[i]])
>   mtext(text= paste(Table4$Exposure[i], "\n", "rho > = ", Table4$rho[i], sep = ""),
+          side = 3)
>  }
```

Figure H.2 shows how, for large values of r, the investigator can increase the number of participants in exchange for reducing the number of repeated measurements without a significant increase of the cost.

I Simulation study for the accuracy of the Wald test approximation

Our calculations are based on the Wald test. The F test or the t-test are more suitable than the Wald test when the covariance matrix needs to be estimated from the data. In the F test, the denominator degrees of freedom need to be estimated from the data in a non trivial way. The same is true for the degrees of freedom needed in the t-test [1]. We consider this fact significantly hinders the methodological development of the problem and therefore we have chosen the Wald test in small samples. A simulation study was performed in order to evaluate the goodness of the approximation of the F-test by the Wald test. In a simulation study, Manor and Zucker showed that the restricted maximum likelihood (REML) approach with the Satterthwaite approximation for the degrees of freedom for the test statistic gave Type I error rates close to the nominal level even in small samples [2]. We assume this good performance also holds in terms of power, and thus compared the results of this approach with the results of the Wald test. We considered the LDD response pattern. Scenarios explored corresponded to all combinations of the values of the parameters: $\theta = 0$; $\rho = 0.2, 0.8$; $\rho_e = 0.2, 0.7$; $p_e = 0.2, 0.7$; $\pi_M = 0, 0.2$; $N = 10, 20, 30, 50, 200$; and $r = 1, 8$. The remaining parameters were fixed at the values: $\sigma^2 = 4$, $\alpha = 0.05$, $\beta_0 = 1$, $\beta_{\text{time}} = -0.5$ and $\beta_{E_0} = -0.5$.

The simulation study was performed following these steps:

1. For each scenario, the design matrix was simulated according to the values of the parameters ρ_e, p_e, π_M, N and r.

2. The value of the β of interest was obtained according to our formulas (based on the Wald test) in order to achieve a power of 0.8.

3. Multivariate normal response was simulated according to the simulated linear predictor based on the model.

4. The model was fitted to the simulated data and the p-values for both the F-test with the Satterthwaite approximation for the degrees of freedom and the Wald test for the significance of the β of interest were stored.

5. Steps 1 to 4 were iterated 500 times.

6. Both F-test with the Satterthwaite approximation and the Wald test empirical power were obtained as the fraction of times that p-value $\leqslant \alpha$.

7. A p-value for the significance of the difference in the empirical power between the F-test and the Wald test was computed.

A total of 147 scenarios were explored. Results are shown in Figure I.1. For N from about 30 almost no impact of avoiding the Satterthwaite correction were found. For lower values of N, the Wald approximation overestimates the power.

J Simulation study for the impact of an unbalanced design

J.1 Simulations

Our methods allow for designs unbalanced with respect to exposure (through the parameter p_e) and unbalanced with respect to the number of time points (through modeling dropout). In addition, we now examined the effect of subjects not having exactly the same vector of times through a simulation study. In this study, we simulated individual time points for the measurements, not at equidistant points $t_j = j/r, j = 0, 1, \ldots, r$, with r the number of repeated measurements, but at points t'_j normally distributed around t_j. The dispersion of the simulated time points t'_j was controlled by the parameter t_{95} which is the ratio between the length of the 95% confidence interval for t'_j and the gap between two consecutive time points in the balanced design, $1/r$. For instance, $t_{95} = 0$ correspond to a balanced design (i.e., $t'_j = t_j$) $t_{95} = 3$ means that 95% of individuals are measured at j-th measurement in a time interval of length $3/r$ around $t_j = j/r$. An usual situation is when $t_{95} = 1$ which corresponds to the case when 95% of participants have been measured at j-th measurement before any of them have been measured at $(j + 1)$-th measurement. In general, in this simulation study each participant has a different follow-up duration.

For each model, CMD and LDD, 36 scenarios were explored. These scenarios correspond to all combinations of the values of the parameters N = 50, 100, 200 and 500; r = 1, 4 and 8; and t_{95} = 0.01, 1 and 3. The remaining parameters were fixed to the following values: $\rho = 0.5$, $\sigma^2 = 1$, $\rho_e = 0.5$, $p_e = 0.3$, $\alpha = 0.05$. In each scenario, the β of interest was obtained as that value for which our formulas provide a power value of 0.9. Then, the remaining coefficients in the model were fixed at 0.5β. In each scenario, the empirical power was assessed through 1000 simulations.

J.2 Results

Results, summarized in Figure J.1, showed no impact of an unbalanced design under the model (1) for $t_{95} \leqslant 1$ while the empirical power decayed at around 0.85 for $t_{95} = 3$ when the expected value under our formulas was 0.9. For the model (2), the empirical power was around 0.93 for $t_{95} = 1$ and around 0.95 for $t_{95} = 3$.

References

[1] Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data (1st edn.). *Springer Series in Statistics*, vol. 22. Springer-Verlag New York Inc., New York, 2000.

[2] Manor O, Zucker DM. Small sample inference for the fixed effects in the mixed linear model. *Computational Statistics & Data Analysis* 2004; **46**(4):801–817. DOI: 10.1016/j.csda.2003.10.005.

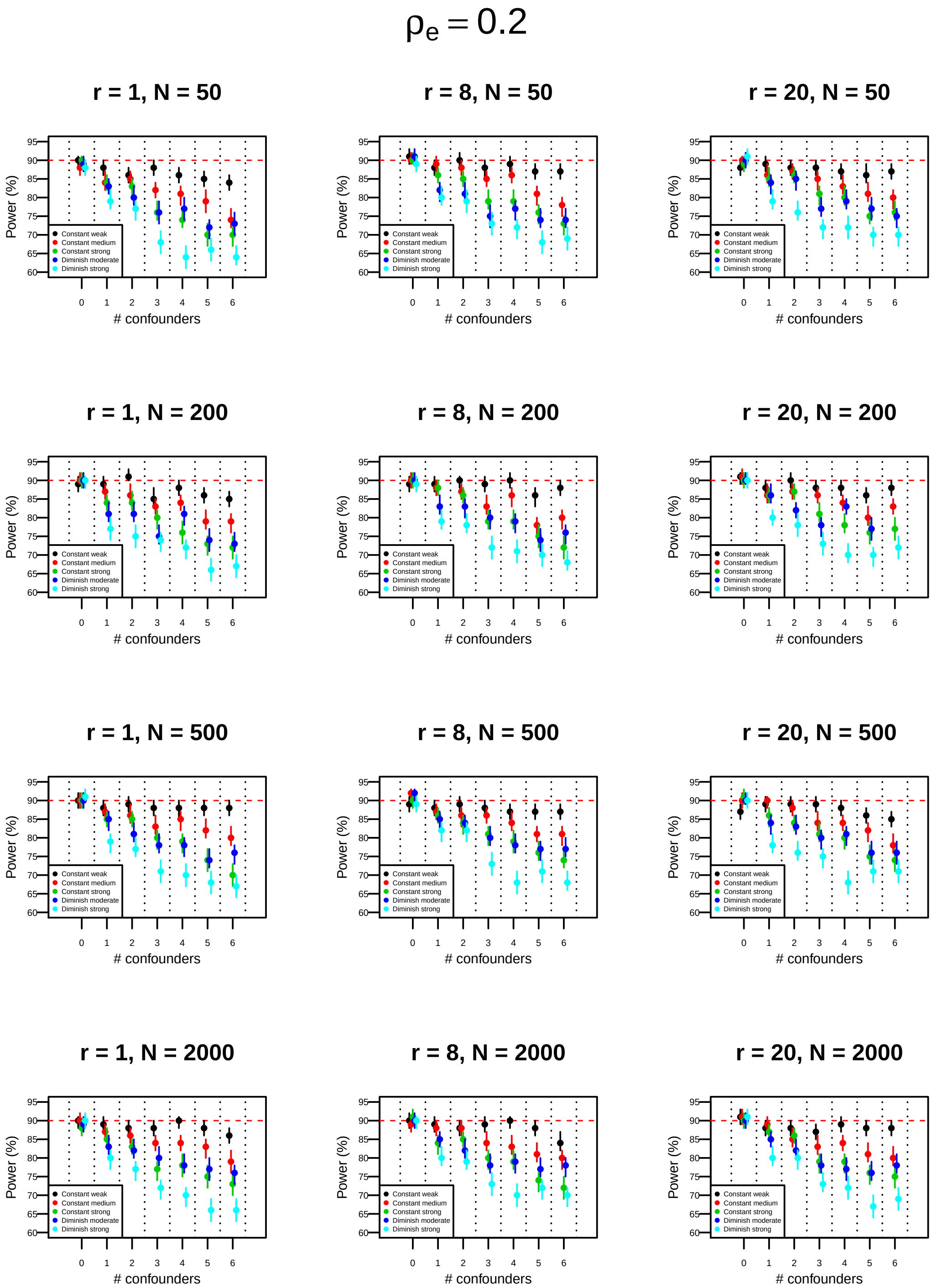

Figure G.1: Impact of confounding under CMD response pattern in scenarios with $\rho_e = 0.2$. For each confounding scenario, empirical mean and percentiles 2.5 and 97.5 were assessed using 1000 simulations.

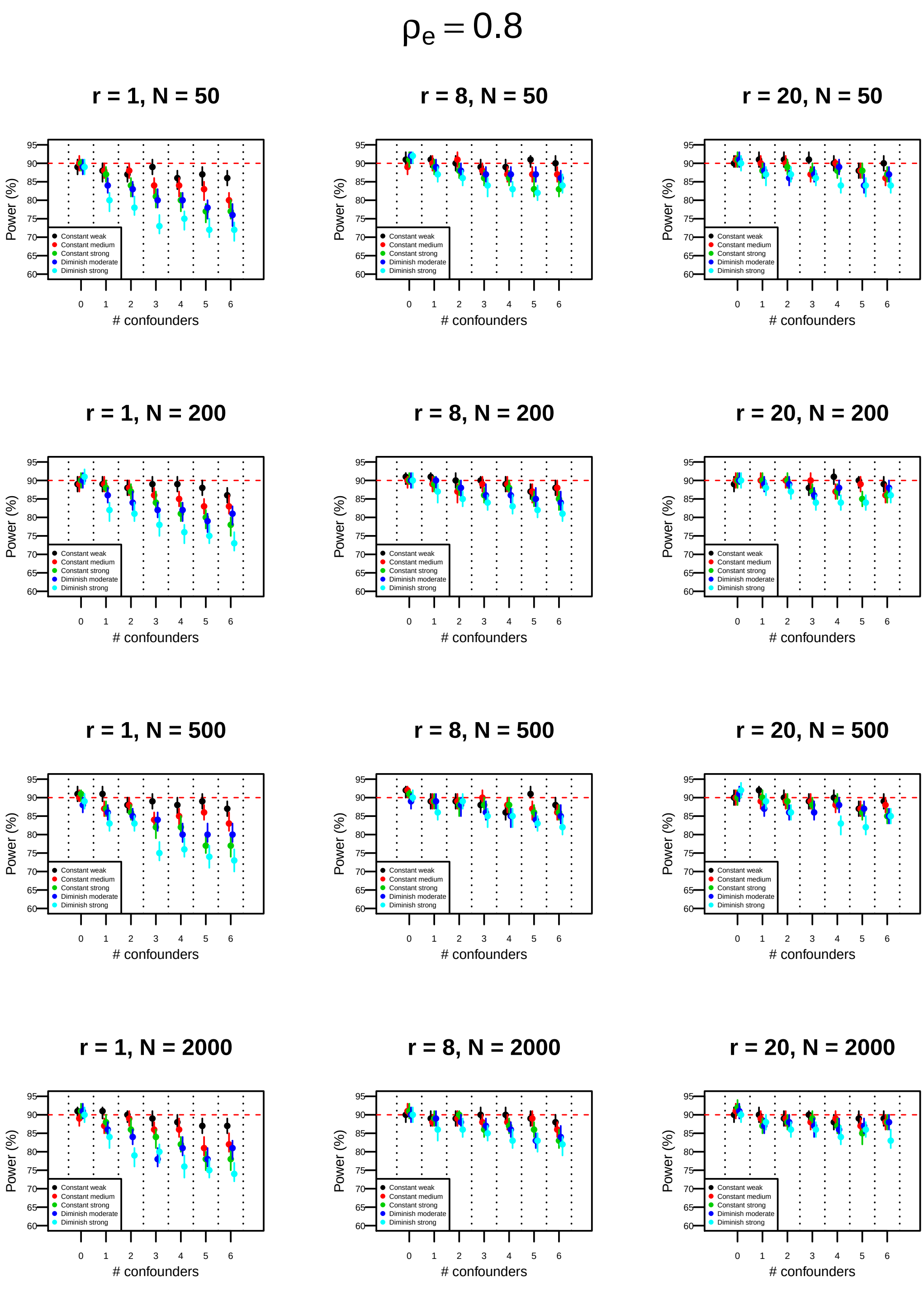

Figure G.2: Impact of confounding under CMD response pattern in scenarios with $\rho_e = 0.8$. For each confounding scenario, empirical mean and percentiles 2.5 and 97.5 were assessed using 1000 simulations.

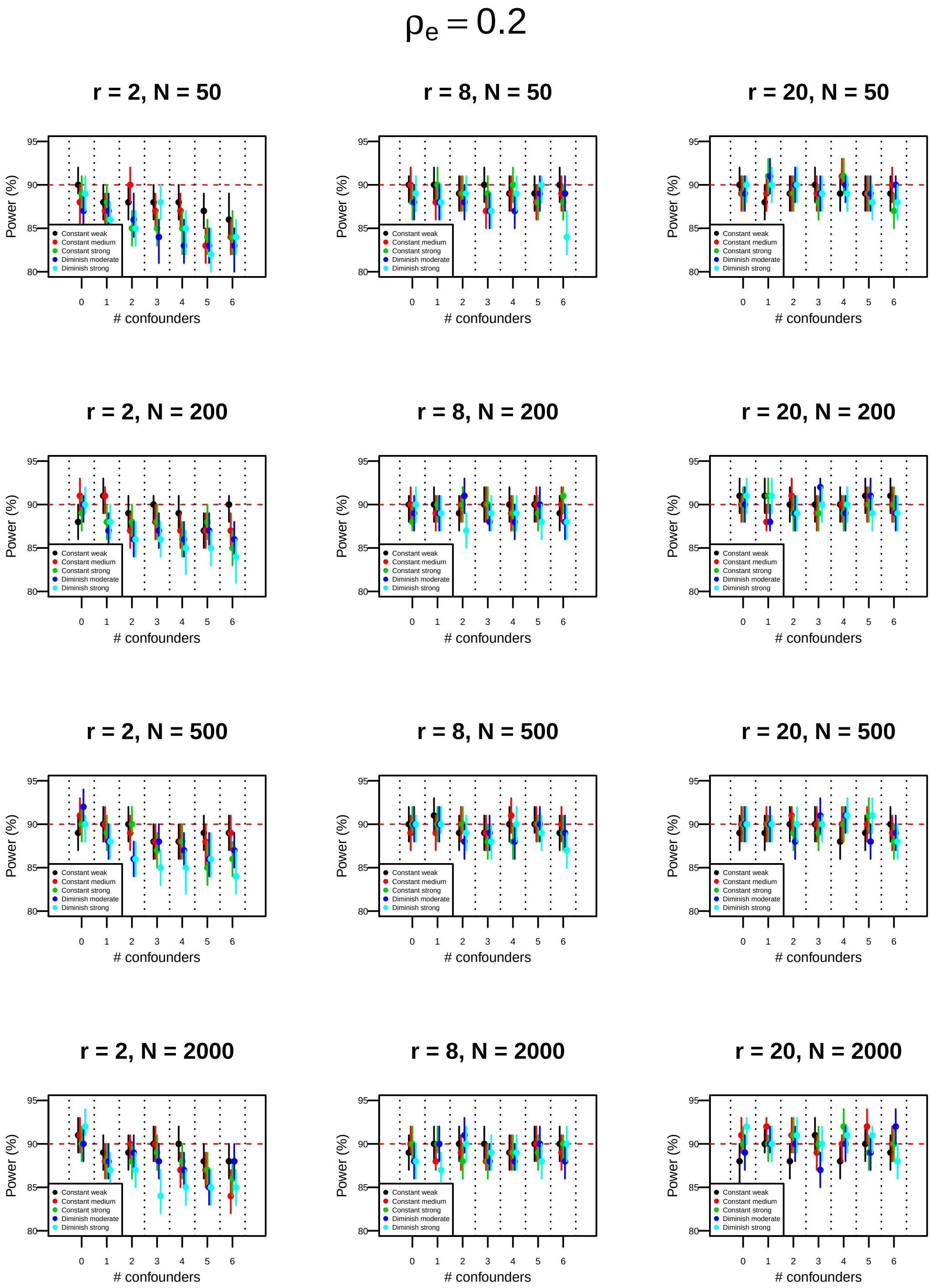

Figure G.3: Impact of confounding under LDD response pattern in scenarios with $\rho_e = 0.2$. For each confounding scenario, empirical mean and percentiles 2.5 and 97.5 were assessed using 1000 simulations.

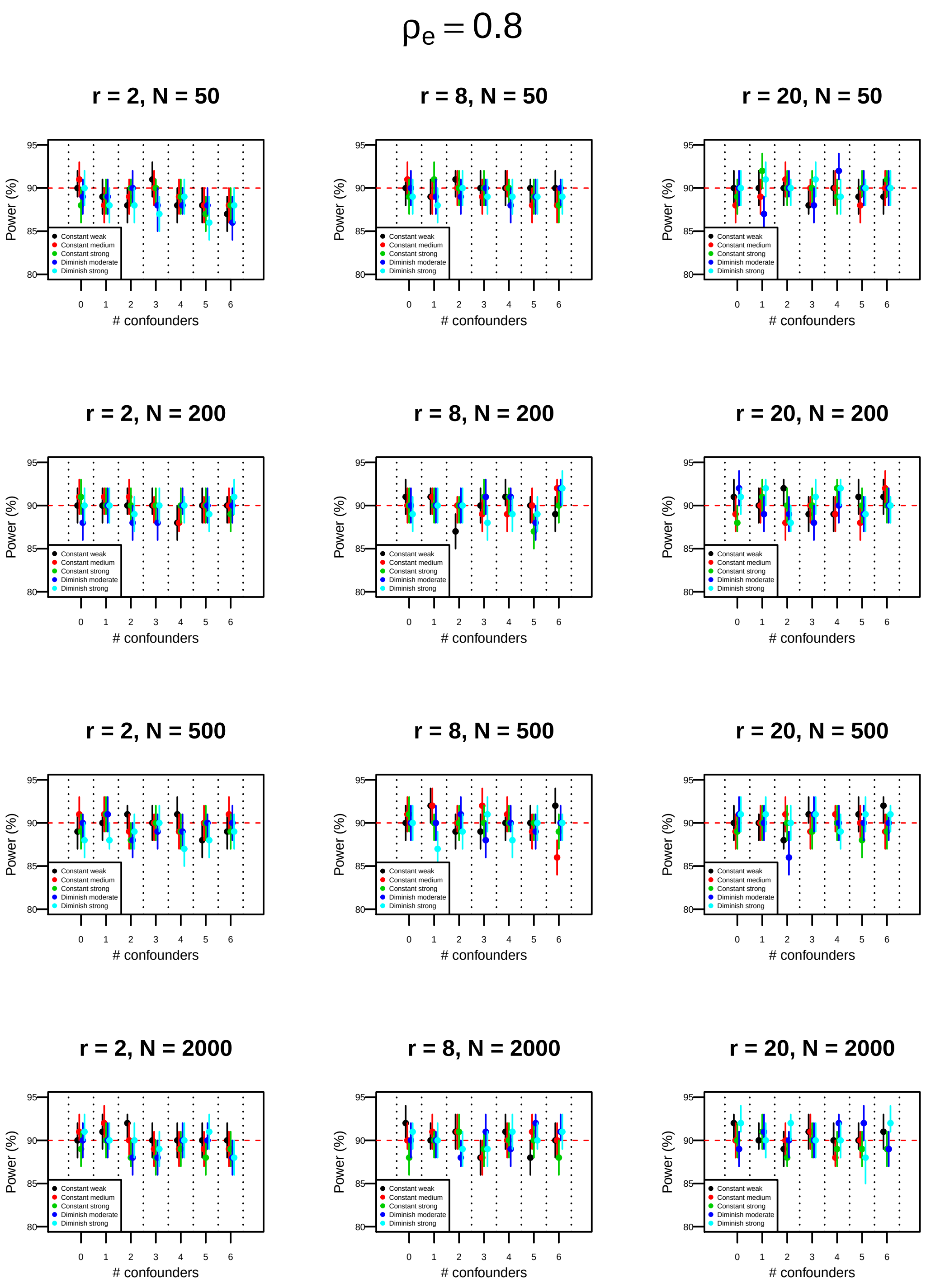

Figure G.4: Impact of confounding under LDD response pattern in scenarios with $\rho_e = 0.8$. For each confounding scenario, empirical mean and percentiles 2.5 and 97.5 were assessed using 1000 simulations.

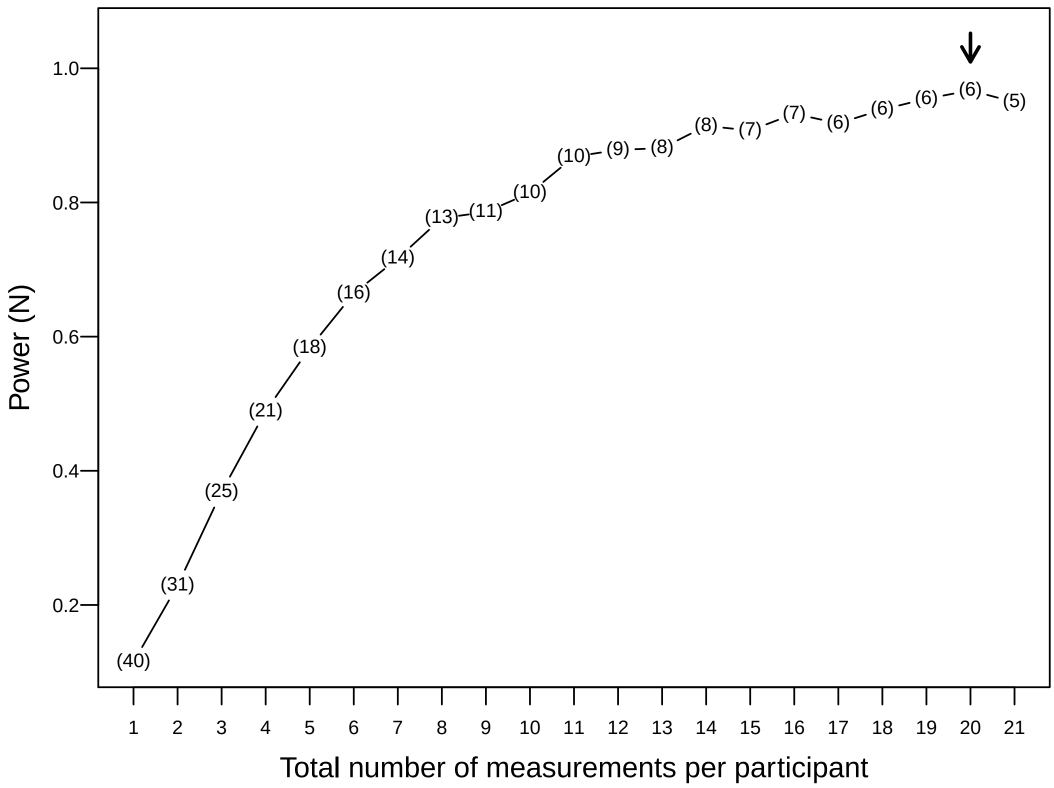

Figure H.1: Maximized power and number of participants (in brackets) as a function of the total number of measurements per participant. The arrow points to the optimal allocation.

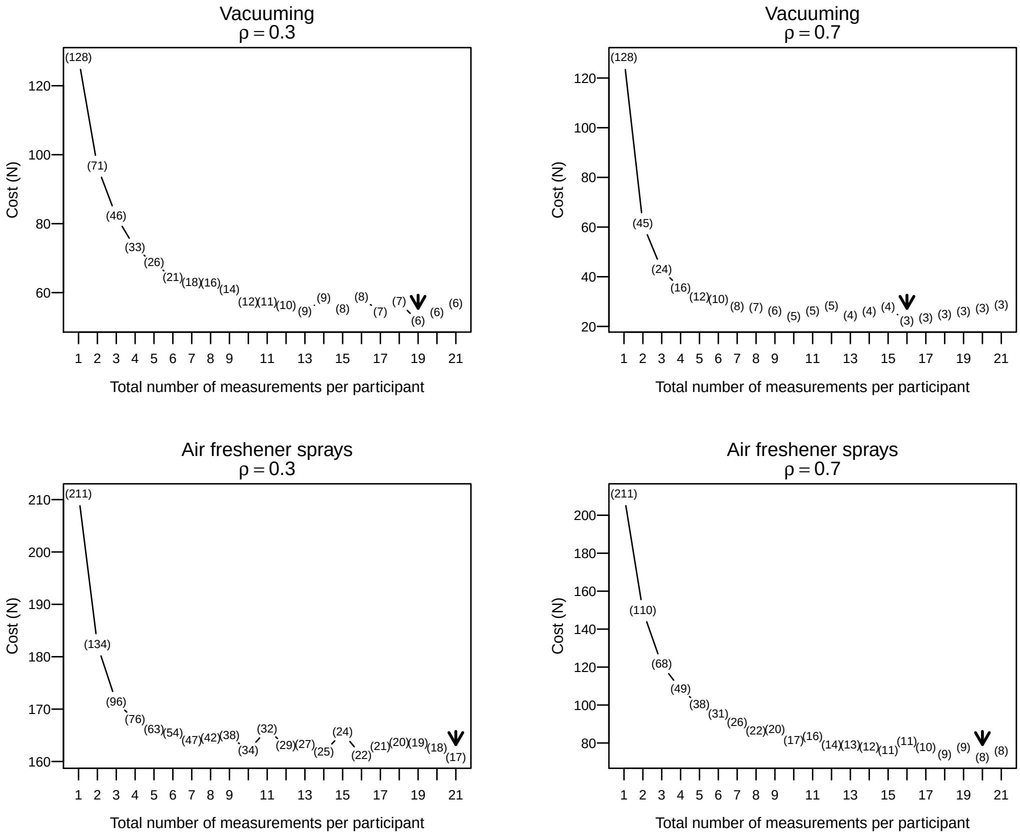

Figure H.2: Minimized cost and number of participants (in brackets) as a function of the total number of measurements per participant. The arrow points to the optimal allocation.

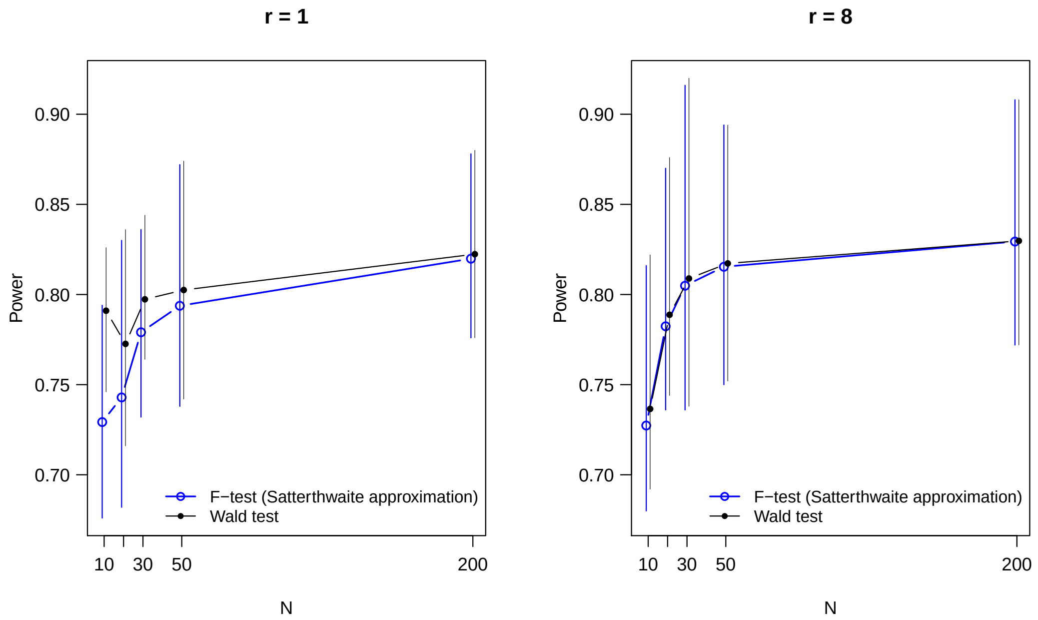

Figure I.1: Empirical power under the F-test with Satterthwaite approximation for the degrees of freedom and the Wald test, for the LDD response pattern. In each of the 147 scenarios, 500 simulations were performed. Points (segments) in the plot area correspond to the mean (minimum and maximum) empirical power between scenarios. In all scenarios, the expected power under our formulas was 0.8.

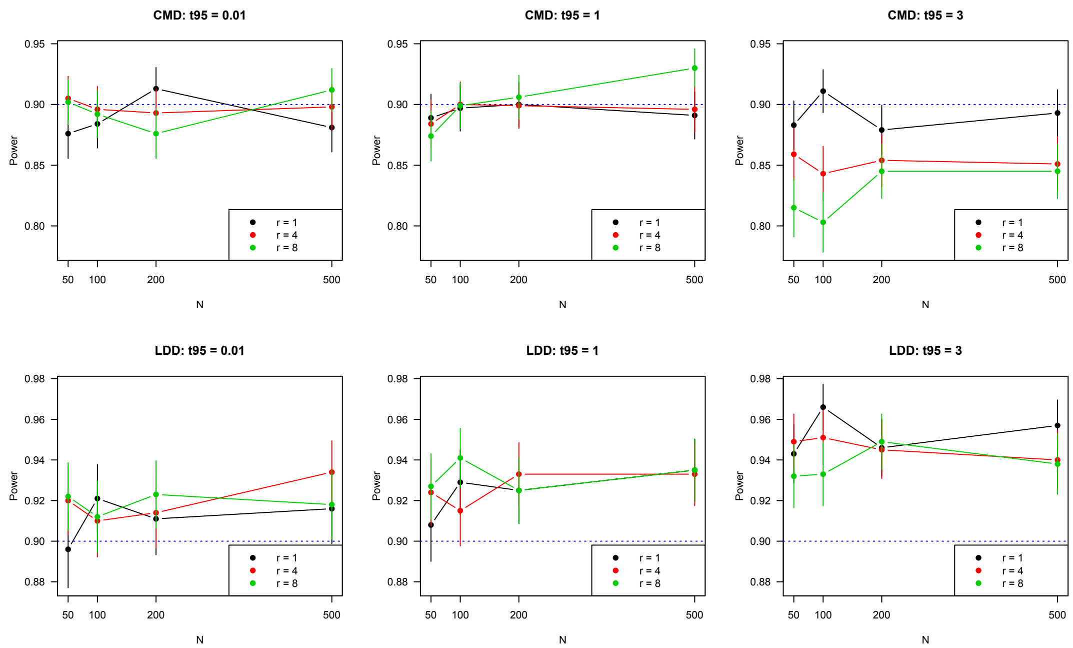

Figure J.1: Results of the simulation study for the impact of an unbalanced design.