Supporting Information for "Relative efficiency of using summary versus individual data in random-effects meta-analysis" by Ding-Geng Chen, Dungang Liu, Xiaoyi Min and Heping Zhang

A Conditions for the Laplace approximation

For simplicity, we denote $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\eta}^\top)^\top$ and $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k^\top, \boldsymbol{\eta}_k^\top)^\top$, let p_k be the dimension and $\boldsymbol{\Theta}_k$ be the domain of $\boldsymbol{\theta}_k$.

(C1) The parameter $\boldsymbol{\theta}_k$ is identifiable. In other words, for any $\boldsymbol{\theta}_k \neq \boldsymbol{\theta}'_k$ in $\boldsymbol{\Theta}_k$, $f_k(y_k, \boldsymbol{x}_k; \boldsymbol{\theta}_k) \neq f_k(y_k, \boldsymbol{x}_k; \boldsymbol{\theta}'_k)$ for some $(y_k, \boldsymbol{x}_k)$;

(C2) For all $(y_k, \boldsymbol{x}_k)$, the function $f_k(y_k, \boldsymbol{x}_k; \boldsymbol{\theta}_k)$ is a three times continuous differentiable function of $\boldsymbol{\theta}_k$ and is positive for all $\boldsymbol{\theta}_k$;

(C3) For all $\boldsymbol{\theta}_k^* \in \boldsymbol{\Theta}_k$, there exist a neighborhood $\mathcal{N}_1(\boldsymbol{\theta}_k^*)$, a positive number n_k, and a random variable Z_1 such that $E_{\boldsymbol{\theta}_k^*}(Z_1) < \infty$ and for all $\boldsymbol{\theta}_k \in \mathcal{N}_1(\boldsymbol{\theta}_k^*)$,

$$\frac{1}{n_k} \sum_{i=1}^{n_k} \log \frac{f_k(y_{ki}, \boldsymbol{x}_{ki}; \boldsymbol{\theta}_k)}{f_k(y_{ki}, \boldsymbol{x}_{ki}; \boldsymbol{\theta}_k^*)} < Z_1;$$

(C4) For all $\boldsymbol{\theta}_k^* \in \boldsymbol{\Theta}_k$, there exist a neighborhood $\mathcal{N}_2(\boldsymbol{\theta}_k^*)$ and a random variable Z_2 such that $E_{\boldsymbol{\theta}_k^*}(Z_2) < \infty$ and for all $\boldsymbol{\theta}_k \in \mathcal{N}_2(\boldsymbol{\theta}_k^*)$, all $1 \leq d \leq 3$, and all $1 \leq j_i, \ldots, j_d \leq p_k$, these is

$$\left| \frac{\partial^d \log f_k(y_k, \boldsymbol{x}_k; \boldsymbol{\theta}_k)}{\partial \theta_{kj_1} \cdots \partial \theta_{kj_d}} \right| < Z_2;$$

(C5) For all $\boldsymbol{\theta}_k^* \in \boldsymbol{\Theta}_k$, define M to be the Hessian matrix of $E_{\boldsymbol{\theta}_k^*}[\log f_k(Y_k, \boldsymbol{X}_k; \boldsymbol{\theta}_k^*) - \log f_k(Y_k, \boldsymbol{X}_k; \boldsymbol{\theta}_k)]$, then $\det(M) > 0$;

(C6) For all $\boldsymbol{\theta}_k^* \in \boldsymbol{\Theta}_k$, the maximum likelihood estimate is strongly consistent.

(C7) $\boldsymbol{\theta}$ lies in the interior of a compact set within the parameter space.

B Proofs

Proof of Lemma 1. According to Theorems 7 and 8 in Kass et al. (1990), under conditions (C1)-(C6) in Appendix A, the sequence of log-likelihood functions $\{\ell_k(\boldsymbol{\beta}_k, \boldsymbol{\eta}_k), n_k = 1, \ldots\}$ is "Laplace regular" with probability one for any true value $\boldsymbol{\theta}_k^*$ in $\boldsymbol{\Theta}$. The proof of Lemma 1 is then similar to Theorem 1 of Kass et al. (1990). In particular, we only need to consider the integration in equation (2.4) over $B_\delta(\hat{\boldsymbol{\theta}}_k) \subseteq \boldsymbol{\Theta}$ for any $0 < \delta < \delta_0$, where $B_\delta(\hat{\boldsymbol{\theta}}_k)$ is the open ball of radius δ centered at $\hat{\boldsymbol{\theta}}_k$. In equation (2.4), we expand $\ell(\boldsymbol{\theta})$ around $\hat{\boldsymbol{\theta}}_k$ to the third order and keep the other terms unchanged. With arguments similar to those in Kass et al. (1990), the third order expansion from $\ell(\boldsymbol{\theta})$ leads to the following leading term of (2.4) and an error term of $O(n_k^{-1/2})$-order:

$$\begin{aligned}
&\ell_k(\hat{\boldsymbol{\theta}}_k) - \frac{1}{2}\log|\boldsymbol{\Sigma}| + \log\int_{B_\delta(\hat{\boldsymbol{\theta}}_k)} \exp[-\frac{1}{2}(\boldsymbol{\theta}_k - \hat{\boldsymbol{\theta}}_k)^\top \mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)(\boldsymbol{\theta}_k - \hat{\boldsymbol{\theta}}_k) - \frac{1}{2}(\boldsymbol{\theta}_k - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_k - \boldsymbol{\theta})]d\boldsymbol{\theta}_k \\
&= \log\int_{B_\delta(\hat{\boldsymbol{\theta}}_k)} \exp\{-\frac{1}{2}(\boldsymbol{\theta}_k - \tilde{\boldsymbol{\theta}}_k)^\top[\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k) + \boldsymbol{\Sigma}^{-1}](\boldsymbol{\theta}_k - \tilde{\boldsymbol{\theta}}_k) - \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k)^\top[\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k)\}d\boldsymbol{\theta}_k \\
&+ \ell_k(\hat{\boldsymbol{\theta}}_k) - \frac{1}{2}\log|\boldsymbol{\Sigma}| \\
&\approx -\frac{1}{2}\log|\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k) + \boldsymbol{\Sigma}^{-1}| - \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k)^\top[\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k) + \ell_k(\hat{\boldsymbol{\theta}}_k) - \frac{1}{2}\log|\boldsymbol{\Sigma}|
\end{aligned} \tag{B.2}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{T} & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}^\top & \boldsymbol{\Phi} \end{pmatrix},$$

and

$$\tilde{\boldsymbol{\theta}}_k = [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k) + \boldsymbol{\Sigma}^{-1}]^{-1}[\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)\hat{\boldsymbol{\theta}}_k + \boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}].$$

Since $\tilde{\boldsymbol{\theta}}_k \in B_\delta(\hat{\boldsymbol{\theta}}_k)$ when n_k is large enough, we can expand the range of integration to the whole parameter space in (B.2) as Kass et al. (1990), which leads to the approximation with an error of exponential decreasing order. Lemma 1 is proved. □

Proof of Lemma 2. Denote the leading term in approximation (2.6) as $\tilde{\ell}_k(\boldsymbol{\theta}, \boldsymbol{T}, \boldsymbol{\Phi}, \boldsymbol{\Xi})$, then

$$\begin{aligned}
& \sum_{k=1}^K \tilde{\ell}_k(\boldsymbol{\theta}, \boldsymbol{T}, \boldsymbol{\Phi}, \boldsymbol{\Xi}) \\
= \; & -\frac{1}{2}\sum_{k=1}^K (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k)^\top [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k) - \frac{1}{2}\sum_{k=1}^K \log|\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k) + \boldsymbol{\Sigma}^{-1}| + C \\
= \; & -\frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top \{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) - \frac{1}{2}\sum_{k=1}^K \hat{\boldsymbol{\theta}}_k^\top [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\hat{\boldsymbol{\theta}}_k \\
+ \; & \frac{1}{2}\{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\hat{\boldsymbol{\theta}}_k\}^\top \{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\}^{-1}\{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\hat{\boldsymbol{\theta}}_k\} \\
- \; & \frac{1}{2}\sum_{k=1}^K \log|\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k) + \boldsymbol{\Sigma}^{-1}| + C,
\end{aligned} \tag{B.3}$$

where $\tilde{\boldsymbol{\theta}} = \{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\}^{-1}\{\sum_{k=1}^K [\mathcal{I}_k(\hat{\boldsymbol{\theta}}_k)^{-1} + \boldsymbol{\Sigma}]^{-1}\hat{\boldsymbol{\theta}}_k\}$. Thus, the leading term in (2.7) maximizes $\sum_{k=1}^K \tilde{\ell}_k(\boldsymbol{\theta}, \boldsymbol{T}, \boldsymbol{\Phi}, \boldsymbol{\Xi})$. (2.7) can be proved by noticing that for any $\epsilon > 0$, for any sequence $\{\boldsymbol{\theta}^{(n)}\}$ such that $||\boldsymbol{\theta}^{(n)} - \tilde{\boldsymbol{\theta}}^{(n)}|| > \epsilon$, $\sum_{k=1}^K \ell_k(\tilde{\boldsymbol{\theta}}^{(n)}, \boldsymbol{T}, \boldsymbol{\Phi}, \boldsymbol{\Xi}) - \sum_{k=1}^K \ell_k(\boldsymbol{\theta}^{(n)}, \boldsymbol{T}, \boldsymbol{\Phi}, \boldsymbol{\Xi})$ is greater than 0 with probability approaching 1. □

Proof of Theorem 1. To prove Part (a), it suffices to show that

$$\widehat{\text{var}}\left(\hat{\boldsymbol{\beta}}_{SS}\right) \geq \widehat{\text{var}}\left(\hat{\boldsymbol{\beta}}_{IPD}\right), \tag{B.4}$$

i.e., the matrix $\widehat{\text{var}}\left(\hat{\boldsymbol{\beta}}_{SS}\right) - \widehat{\text{var}}\left(\hat{\boldsymbol{\beta}}_{IPD}\right)$ is positive semi-definite. After some algebraic calculations, we can show that

$$\widehat{\text{var}}\left(\hat{\boldsymbol{\beta}}_{SS}\right) = \left(\sum_{k=1}^K A_k^{-1}\right)^{-1}$$

and

$$\widehat{\operatorname{var}}\left(\hat{\boldsymbol{\beta}}_{IPD}\right) = \left\{ \sum_{k=1}^{K} \begin{pmatrix} A_k & B_k \\ B_k^\top & C_k \end{pmatrix}^{-1} \right\}^{-1}_{[A_k]},$$

where

$$A_k = \left(\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k} - \mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k}\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k}\right)^{-1}_{|\hat{\boldsymbol{\beta}}_k,\hat{\boldsymbol{\eta}}_k} + \hat{\boldsymbol{T}},$$

$$B_k = -\mathcal{I}^{-1}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k}\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\left(\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k} - \mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k}\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\right)^{-1}_{|\hat{\boldsymbol{\beta}}_k,\hat{\boldsymbol{\eta}}_k} + \hat{\boldsymbol{\Xi}},$$

$$C_k = \left(\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k} - \mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k}\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\right)^{-1}_{|\hat{\boldsymbol{\beta}}_k,\hat{\boldsymbol{\eta}}_k} + \hat{\boldsymbol{\Phi}}.$$

If we let

$$\begin{pmatrix} A_k & B_k \\ B_k^\top & C_k \end{pmatrix}^{-1} = \begin{pmatrix} \tilde{A}_k & \tilde{B}_k \\ \tilde{B}_k^\top & \tilde{C}_k \end{pmatrix},$$

then the inequality (B.4) is equivalent to

$$\left\{ \sum_{k=1}^{K} \left(\tilde{A}_k - \tilde{B}_k\tilde{C}_k^{-1}\tilde{B}_k^\top\right) \right\}^{-1} \geq \left\{ \sum_{k=1}^{K} \tilde{A}_k - \sum_{k=1}^{K} \tilde{B}_k \left(\sum_{k=1}^{K} \tilde{C}_k\right)^{-1} \sum_{k=1}^{K} \tilde{B}_k^\top \right\}^{-1}.$$

The above inequality could be further simplified as

$$\sum_{k=1}^{K} \tilde{B}_k\tilde{C}_k^{-1}\tilde{B}_k^\top \geq \sum_{k=1}^{K} \tilde{B}_k \left(\sum_{k=1}^{K} \tilde{C}_k\right)^{-1} \sum_{k=1}^{K} \tilde{B}_k^\top,$$

which holds according to Lemma 1 in the Appendix A of Lin and Zeng (2010). This in turn establishes the inequality (B.4).

The proof of the inequality in Part (b) is similar to that of Part (a) except that we use the true values of the parameters $\boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{T}, \boldsymbol{\Xi}, \boldsymbol{\Phi}$ to evaluate A_k, B_k, C_k. The equality is achieved if and only if

$$\tilde{B}_1\tilde{C}_1^{-1} = \tilde{B}_2\tilde{C}_2^{-1} = \cdots = \tilde{B}_K\tilde{C}_K^{-1}.$$

Since $\tilde{B}_k = -A_k^{-1}B_k(C_k - B_k^\top A_k^{-1}B_k)^{-1}$ and $\tilde{C}_k^{-1} = C_k - B_k^\top A_k^{-1}B_k$, the above condition is equivalent to

$$A_1^{-1}B_1 = A_2^{-1}B_2 = \cdots = A_K^{-1}B_K.$$

Therefore, given $\boldsymbol{\Xi} = \mathbf{0}$, the equality can be achieved only if the quantity

$$\left\{ \left(\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k} - \mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k}\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k} \right)^{-1} + \boldsymbol{T} \right\}^{-1} \mathcal{I}^{-1}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k}\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k} \left(\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k} - \mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k}\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k} \right)^{-1}$$

does not depend on the study index k. This condition will not hold except in special cases. Therefore, the inequality may hold even if $\boldsymbol{\Xi} = \mathbf{0}$.

Following the arguments in the proof of Part (b), the equality in Part (c) holds. To see this, notice that $A_i^{-1}B_i \to \boldsymbol{T}^{-1}\boldsymbol{\Xi}$ as the sample size in each study $n \to \infty$. Moreover, $|A_i^{-1}B_i - A_j^{-1}B_j| = O(1/n)$ for any $i \neq j$. Thus, as $n \to \infty$, $K \to \infty$ and $Kn^{-1/2} \to 0$, $\sum_{i \neq j} |A_i^{-1}B_i - A_j^{-1}B_j| \to 0$. This guarantees that the equality is achieved in the limit. $\square$

C A special case of Section 2.3

We consider a special case where the likelihood function in (2.5) is exact in the sense that there is no approximation error term $O(n_k^{-\frac{1}{2}})$. Specifically, assume that in the k-th study, the log-likelihood function is of the following form

$$\ell_k(\boldsymbol{\beta}_k, \boldsymbol{\eta}_k) = -\begin{pmatrix} \boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_k \\ \boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k \end{pmatrix}^{\top} \begin{pmatrix} \mathcal{I}_{k,11} & \mathcal{I}_{k,12} \\ \mathcal{I}_{k,21} & \mathcal{I}_{k,22} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_k \\ \boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k \end{pmatrix} + c_k,$$

where c_k is a statistic that does not depend on any parameters. For clarity, we assume that all the variance components, including the within-study covariance $\mathcal{I}_k$ and the between-study variance $\boldsymbol{T}$, are known.

For this special case, the summary statistic

$$\hat{\boldsymbol{\beta}}_k \sim N\left(\boldsymbol{\beta}_k, \mathrm{var}_C(\hat{\boldsymbol{\beta}}_k) + \boldsymbol{T} \right),$$

where the exact variance $\mathrm{var}_C(\hat{\boldsymbol{\beta}}_k) = \left(\mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\beta}_k} - \mathcal{I}_{k,\boldsymbol{\beta}_k\boldsymbol{\eta}_k}\mathcal{I}^{-1}_{k,\boldsymbol{\eta}_k\boldsymbol{\eta}_k}\mathcal{I}_{k,\boldsymbol{\eta}_k\boldsymbol{\beta}_k} \right)^{-1}$. The combined estimator is

$$\hat{\boldsymbol{\beta}}_{SS} = \left[\sum_{k=1}^{K}\left\{\text{var}_C(\hat{\boldsymbol{\beta}}_k) + \boldsymbol{T}\right\}^{-1}\right]^{-1}\sum_{k=1}^{K}\left\{\text{var}_C(\hat{\boldsymbol{\beta}}_k) + \boldsymbol{T}\right\}^{-1}\hat{\boldsymbol{\beta}}_k.$$

The exact variance of $\hat{\boldsymbol{\beta}}_{SS}$ is

$$\text{var}(\hat{\boldsymbol{\beta}}_{SS}) = \left[\sum_{k=1}^{K}\left\{\left(\mathcal{I}_{k,11} - \mathcal{I}_{k,12}\mathcal{I}_{k,22}^{-1}\mathcal{I}_{k,21}\right)^{-1} + \boldsymbol{T}\right\}^{-1}\right]^{-1}.$$

To compare $\text{var}(\hat{\boldsymbol{\beta}}_{SS})$ with $\text{var}(\hat{\boldsymbol{\beta}}_{IPD})$, we derive the log-likelihood function of $(\boldsymbol{\beta}, \boldsymbol{\eta})$ in (2.6), which simplifies to

$$\ell_k(\boldsymbol{\beta}, \boldsymbol{\eta}) = -\begin{pmatrix}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_k \\ \boldsymbol{\eta} - \hat{\boldsymbol{\eta}}_k\end{pmatrix}^{\top}\left[\begin{pmatrix}\mathcal{I}_{k,11} & \mathcal{I}_{k,12} \\ \mathcal{I}_{k,21} & \mathcal{I}_{k,22}\end{pmatrix}^{-1} + \begin{pmatrix}\boldsymbol{T} & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}^{\top} & \boldsymbol{\Phi}\end{pmatrix}\right]^{-1}\begin{pmatrix}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_k \\ \boldsymbol{\eta} - \hat{\boldsymbol{\eta}}_k\end{pmatrix} + c_k.$$

Therefore, the IPD estimator is

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}_{IPD} \\ \hat{\boldsymbol{\eta}}_{IPD}\end{pmatrix} = \left\{\left[\sum_{k=1}^{K}\mathcal{M}_k^{-1}\right]^{-1}\left[\sum_{k=1}^{K}\mathcal{M}_k^{-1}\begin{pmatrix}\hat{\boldsymbol{\beta}}_k \\ \hat{\boldsymbol{\eta}}_k\end{pmatrix}\right]\right\},$$

where

$$\mathcal{M}_k = \begin{pmatrix}\mathcal{I}_{k,11} & \mathcal{I}_{k,12} \\ \mathcal{I}_{k,21} & \mathcal{I}_{k,22}\end{pmatrix}^{-1} + \begin{pmatrix}\boldsymbol{T} & \boldsymbol{\Xi} \\ \boldsymbol{\Xi}^{\top} & \boldsymbol{\Phi}\end{pmatrix}.$$

The exact variance of $\hat{\boldsymbol{\beta}}_{IPD}$ is

$$\text{var}\left(\hat{\boldsymbol{\beta}}_{IPD}\right) = \left\{\sum_{k=1}^{K}\mathcal{M}_k^{-1}\right\}^{-1}_{[1,1]}.$$

Similar to the proof of Theorem 1, we can show that

$$\text{var}(\hat{\boldsymbol{\beta}}_{SS}) \geq \text{var}\left(\hat{\boldsymbol{\beta}}_{IPD}\right).$$

The equality can be achieved if (a) the within-study correlation $\mathcal{I}_{k,12} = \mathbf{0}$ and the between-study correlation $\boldsymbol{\Xi} = \mathbf{0}$; or (b) $\mathcal{M}_k$ is the same for all k's. The above inequality is established provided that all variance components are known. Given estimates of variance components, the inequality still holds in the settings considered in our simulation studies.

The result in this subsection indicates that we can establish an inequality between the exact variances of $\hat{\boldsymbol{\beta}}_{SS}$ and $\hat{\boldsymbol{\beta}}_{IPD}$, when each individual log-likelihood function can be written as a quadratic form of $(\boldsymbol{\beta}_k, \boldsymbol{\eta}_k)$ without an approximation error. For the general case where approximation errors may exist, the inequalities in Theorem 1(a)-(b) are established between the variance estimates or between the asymptotic variances.

D Supplementary tables for simulation studies

Table 6: The IPD- and summary-statistics-based meta-analysis estimates of the interaction effect β_3 for continuous outcomes (with estimated variance components $\sigma_\epsilon = 10$ and Σ_β).

		IPD					Summary statistics				
K	n_k	Mean	SE	ESE	$\sqrt{\text{MSE}}$	CP	Mean	SE	ESE	$\sqrt{\text{MSE}}$	CP
5	20	1.058	2.184	2.295	2.184	0.959	1.043	2.393	2.503	2.392	0.952
	50	1.030	1.354	1.440	1.354	0.958	1.030	1.401	1.499	1.400	0.961
	100	1.003	0.885	1.022	0.885	0.971	1.003	0.913	1.044	0.913	0.971
	200	0.963	0.683	0.743	0.684	0.968	0.965	0.689	0.752	0.689	0.965
	500	1.001	0.466	0.488	0.466	0.950	1.000	0.466	0.491	0.466	0.953
10	20	1.021	1.521	1.593	1.520	0.953	0.966	1.699	1.775	1.699	0.959
	50	0.991	0.939	0.998	0.939	0.961	0.990	0.977	1.037	0.976	0.958
	100	1.020	0.653	0.717	0.653	0.969	1.017	0.669	0.732	0.669	0.969
	200	0.991	0.485	0.516	0.485	0.956	0.993	0.490	0.521	0.489	0.953
	500	1.001	0.333	0.343	0.332	0.947	1.001	0.333	0.345	0.333	0.952
30	20	0.984	0.823	0.881	0.823	0.958	1.000	0.946	0.991	0.945	0.954
	50	1.002	0.525	0.560	0.524	0.962	0.996	0.553	0.584	0.552	0.962
	100	1.001	0.391	0.400	0.390	0.944	1.004	0.398	0.409	0.398	0.954
	200	0.994	0.286	0.289	0.285	0.944	0.994	0.288	0.293	0.288	0.949
	500	1.000	0.188	0.194	0.188	0.951	0.999	0.189	0.195	0.189	0.954
50	20	1.008	0.676	0.675	0.676	0.940	1.005	0.772	0.762	0.771	0.944
	50	0.998	0.403	0.429	0.402	0.961	1.009	0.425	0.450	0.425	0.962
	100	0.999	0.296	0.307	0.296	0.954	1.002	0.307	0.315	0.307	0.949
	200	1.007	0.215	0.222	0.215	0.957	1.005	0.214	0.224	0.214	0.955
	500	0.998	0.145	0.148	0.145	0.951	0.999	0.145	0.148	0.145	0.953
100	20	0.983	0.450	0.470	0.451	0.963	0.980	0.525	0.533	0.525	0.953
	50	1.000	0.278	0.299	0.278	0.968	1.003	0.292	0.314	0.292	0.971
	100	1.010	0.218	0.213	0.218	0.941	1.011	0.223	0.219	0.224	0.943
	200	1.002	0.153	0.155	0.153	0.953	1.003	0.155	0.157	0.155	0.949
	500	1.000	0.102	0.104	0.102	0.948	1.000	0.103	0.105	0.103	0.952

Mean– the average of estimates for 1000 simulation replicates; SE–standard error of the estimates from 1000 simulation runs; ESE–estimated standard error of the estimate for each simulation run; MSE–mean squared error; CP–coverage probability of 95% confidence intervals.

Table 7: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for continuous outcomes (with estimated variance components σ_ϵ and Σ_β) when $\rho_{x,k} \sim$ Uniform$(-0.3, 0.3)$.

σ_ϵ	$n_k \setminus K$	5	10	30	50	100
1	20	1.033	1.042	1.065	1.080	1.102
	50	0.994	1.004	1.007	1.013	1.001
	100	1.001	1.005	1.003	1.002	1.005
	200	0.998	1.000	1.000	1.000	0.999
	500	1.000	1.000	1.000	1.000	1.000
3	20	1.142	1.142	1.184	1.180	1.254
	50	1.013	1.020	1.053	1.037	1.044
	100	1.010	1.020	1.013	1.012	1.031
	200	0.997	1.000	1.005	0.999	0.995
	500	0.994	0.997	1.000	1.000	1.004
10	20	1.200	1.249	1.319	1.302	1.360
	50	1.070	1.082	1.110	1.113	1.106
	100	1.064	1.049	1.040	1.075	1.056
	200	1.016	1.018	1.017	0.998	1.025
	500	1.000	1.005	1.015	1.003	1.014
30	20	1.208	1.271	1.345	1.328	1.376
	50	1.077	1.109	1.117	1.132	1.122
	100	1.085	1.069	1.045	1.099	1.070
	200	1.018	1.028	1.030	1.021	1.034
	500	1.027	1.022	1.030	1.026	1.029

Table 8: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for continuous outcomes when $\rho_{x,k} = 0$.

σ_ϵ	$n_k \setminus K$	5	10	30	50	100
1	20	1.051	1.024	1.047	1.023	1.039
	50	1.005	1.001	1.011	1.003	1.010
	100	1.000	1.005	1.001	1.003	1.003
	200	0.999	0.999	1.002	1.000	1.001
	500	1.000	1.000	1.000	1.000	1.000
3	20	1.164	1.134	1.174	1.134	1.141
	50	1.021	1.034	1.053	1.054	1.033
	100	1.007	1.010	0.989	1.017	1.004
	200	1.006	0.998	1.006	1.002	1.001
	500	1.002	1.002	0.998	1.002	1.002
10	20	1.227	1.260	1.288	1.269	1.257
	50	1.047	1.089	1.112	1.109	1.084
	100	1.034	1.022	1.018	1.028	1.024
	200	1.011	1.002	1.014	1.022	1.008
	500	1.002	1.004	0.994	1.013	1.003
30	20	1.229	1.276	1.309	1.303	1.282
	50	1.052	1.093	1.130	1.122	1.105
	100	1.042	1.035	1.036	1.043	1.033
	200	1.019	1.011	1.021	1.030	1.017
	500	1.002	1.010	0.995	1.015	1.006

Table 9: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for continuous outcomes when $\rho_{x,k} \sim$ Uniform$(0, 0.7)$.

σ_ϵ	$n_k \setminus K$	5	10	30	50	100
1	20	1.019	1.027	1.086	1.065	1.088
	50	0.997	1.015	1.014	1.015	1.000
	100	1.002	1.002	1.007	1.007	1.006
	200	1.000	0.998	1.001	0.999	0.999
	500	1.000	0.999	1.000	1.001	1.000
3	20	1.106	1.139	1.278	1.153	1.213
	50	1.031	1.047	1.075	1.056	1.031
	100	1.010	1.012	1.022	1.020	1.038
	200	0.998	0.999	0.998	0.983	1.000
	500	0.992	0.992	1.000	1.000	1.004
10	20	1.198	1.212	1.418	1.260	1.316
	50	1.096	1.094	1.133	1.117	1.096
	100	1.047	1.050	1.061	1.086	1.062
	200	1.010	1.005	1.024	1.005	1.022
	500	1.004	1.013	1.020	1.003	1.011
30	20	1.218	1.230	1.435	1.294	1.334
	50	1.113	1.115	1.145	1.129	1.113
	100	1.075	1.066	1.066	1.108	1.065
	200	1.020	1.022	1.041	1.025	1.034
	500	1.028	1.035	1.045	1.030	1.037

Table 10: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for continuous outcomes when the between-study correlation $\mathbf{\Xi} = \mathbf{0}$ and $\rho_{x,k} \sim \text{Uniform}(-0.3, 0.3)$ (with known variance components $\sigma_\epsilon = 10$ and Σ_β).

σ_ϵ	$n_k \backslash K$	5	10	30	50	100
1	20	1.055	1.061	1.074	1.056	1.075
	50	1.006	1.008	1.006	1.006	1.000
	100	1.000	1.004	1.003	1.001	1.005
	200	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000
3	20	1.192	1.201	1.235	1.204	1.245
	50	1.038	1.043	1.055	1.051	1.048
	100	1.014	1.026	1.019	1.020	1.029
	200	0.998	1.001	1.002	0.997	1.004
	500	1.000	1.001	1.000	1.000	1.003
10	20	1.253	1.300	1.357	1.325	1.364
	50	1.097	1.104	1.115	1.126	1.115
	100	1.074	1.065	1.047	1.078	1.062
	200	1.009	1.017	1.015	1.003	1.023
	500	1.012	1.011	1.012	1.007	1.017
30	20	1.255	1.316	1.371	1.346	1.379
	50	1.115	1.122	1.124	1.141	1.124
	100	1.100	1.078	1.050	1.105	1.071
	200	1.021	1.030	1.028	1.022	1.032
	500	1.040	1.029	1.032	1.030	1.030

Table 11: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for continuous outcomes when the between-study correlation $\mathbf{\Xi} = \mathbf{0}$ and $\rho_{x,k} \sim$ Uniform$(-0.3, 0.3)$ (with estimated variance components σ_ϵ and Σ_β).

σ_ϵ	$n_k \backslash K$	5	10	30	50	100
1	20	1.024	1.036	1.056	1.055	1.071
	50	0.992	1.000	1.005	1.008	1.001
	100	0.998	1.001	1.003	1.001	1.004
	200	1.000	0.999	1.000	1.000	1.001
	500	1.000	1.000	1.000	1.000	1.000
3	20	1.154	1.171	1.187	1.185	1.236
	50	0.999	1.021	1.053	1.038	1.042
	100	1.001	1.007	1.018	1.019	1.029
	200	0.996	0.997	1.000	0.996	1.002
	500	0.994	1.000	0.999	0.999	1.002
10	20	1.200	1.264	1.331	1.311	1.365
	50	1.063	1.088	1.115	1.114	1.109
	100	1.067	1.055	1.039	1.076	1.060
	200	1.014	1.018	1.016	1.001	1.022
	500	1.007	1.002	1.013	1.002	1.011
30	20	1.209	1.274	1.344	1.328	1.376
	50	1.079	1.109	1.118	1.131	1.123
	100	1.085	1.071	1.045	1.098	1.072
	200	1.019	1.031	1.028	1.021	1.034
	500	1.030	1.025	1.029	1.027	1.032

Table 12: The IPD- and summary-statistics-based meta-analysis estimates of the interaction effect β_3 for binary outcomes when 20% of the K studies have sample sizes $3n_k$.

		IPD					Summary statistics				
K	n_k	Mean	SE	ESE	$\sqrt{\text{MSE}}$	CP	Mean	SE	ESE	$\sqrt{\text{MSE}}$	CP
10	100	0.999	0.874	0.947	0.874	0.967	1.012	0.890	0.968	0.889	0.963
	200	0.969	0.652	0.681	0.652	0.958	0.966	0.655	0.691	0.655	0.961
	500	1.018	0.430	0.441	0.430	0.961	1.014	0.425	0.447	0.425	0.961
	1,000	1.031	0.312	0.339	0.314	0.958	1.030	0.319	0.345	0.320	0.956
	2,000	1.013	0.252	0.271	0.252	0.947	1.006	0.247	0.276	0.247	0.949
30	100	0.969	0.500	0.531	0.500	0.966	0.976	0.507	0.542	0.508	0.968
	200	0.989	0.353	0.379	0.353	0.954	0.987	0.358	0.384	0.358	0.962
	500	1.007	0.243	0.253	0.243	0.968	1.002	0.247	0.257	0.247	0.961
	1,000	1.005	0.176	0.194	0.176	0.968	0.998	0.176	0.198	0.176	0.970
	2,000	0.997	0.142	0.161	0.142	0.978	0.992	0.142	0.164	0.142	0.973
50	100	1.016	0.376	0.406	0.376	0.965	1.018	0.385	0.414	0.385	0.963
	200	1.015	0.274	0.290	0.274	0.960	1.012	0.277	0.294	0.277	0.964
	500	1.007	0.184	0.195	0.184	0.962	0.998	0.187	0.198	0.187	0.968
	1,000	1.012	0.140	0.149	0.140	0.964	1.007	0.141	0.152	0.141	0.967
	2,000	1.005	0.110	0.124	0.110	0.962	0.999	0.111	0.126	0.110	0.965
100	100	1.005	0.269	0.283	0.269	0.958	1.006	0.279	0.289	0.279	0.952
	200	1.013	0.196	0.203	0.196	0.962	1.008	0.198	0.205	0.198	0.962
	500	1.005	0.126	0.137	0.126	0.970	0.998	0.126	0.139	0.126	0.972
	1,000	1.000	0.098	0.105	0.098	0.959	0.993	0.098	0.107	0.098	0.958
	2,000	0.997	0.081	0.088	0.081	0.962	0.991	0.080	0.090	0.081	0.966

Mean– the average of estimates for 1000 simulation replicates; SE–standard error of the estimates from 1000 simulation runs; ESE–estimated standard error of the estimate for each simulation run; MSE–mean squared error; CP–coverage probability of 95% confidence intervals.

Table 13: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for binary outcomes when 20% of the K studies have sample sizes $3n_k$.

$n_k \setminus K$	10	30	50	100
100	1.036	1.030	1.051	1.077
200	1.010	1.027	1.026	1.019
500	0.978	1.028	1.025	0.992
$1,000$	1.040	0.993	1.011	1.018
$2,000$	0.960	0.995	1.004	0.994

Table 14: Relative efficiency of $\hat{\beta}_{3,SS}$ versus $\hat{\beta}_{3,IPD}$ for binary outcomes under the same setting of Table 4. But $\hat{\beta}_{3,SS}$ is obtained from a multivariate random-effects meta-analysis using the summary statistics for $(\boldsymbol{\beta}_k, \boldsymbol{\eta}_k)$.

Correlation	$n_k \setminus K$	10	30	50	100
$\rho_x = 0$	100	0.930	0.940	0.944	0.934
	200	0.969	0.962	0.958	0.976
	500	0.988	0.983	0.984	0.993
	$1,000$	0.991	0.986	0.988	0.992
	$2,000$	0.995	0.997	0.993	0.991
$\rho_x \sim$ Unif(-0.3,0.3)	100	0.932	0.929	0.941	0.945
	200	0.974	0.971	0.966	0.956
	500	0.987	0.981	0.983	0.975
	$1,000$	0.992	0.988	0.987	0.989
	$2,000$	0.996	0.993	0.995	0.991

Table 15: The IPD- and summary-statistics-based meta-analysis of the alcohol intervention data excluding two large studies (8a and 8b).

	IPD			Summary Statistics		
Parameter	Estimate	SE	P-value	Estimate	SE	P-value
Linear regression models for $Y^{(1)}$ (the change in the number of drinks)						
β_1	0.065	0.183	0.722	0.012	0.188	0.950
β_2	-0.370	0.049	4.80×10^{-14}	-0.401	0.054	1.06×10^{-13}
β_3	-0.015	0.046	0.744	-0.030	0.052	0.559
Logistic regression models for $Y^{(2)}$ (whether the number of drinks was reduced)						
β_1	0.404	0.223	0.070	0.264	0.235	0.260
β_2	0.538	0.066	2.35×10^{-16}	0.493	0.069	1.02×10^{-12}
β_3	-0.072	0.050	0.153	-0.034	0.055	0.535

E Supplementary example

Beta-blockade was a major drug to reduce mortality after myocardial infarction in the treatment of patients with myocardial infarction. Yusuf et al. (1985) presented an overview of the effectiveness of beta-blockade during and after myocardial infarction from 22 clinical trial centers. Since its publication, it has been widely cited for meta-analysis and other applications. For example, Freemantle et al. (1999) used it for meta-analysis to assess the effectiveness of beta-blockade in short term treatment for acute myocardial infarction and in longer term secondary prevention.

As shown in Table 16, each center reported the number of deaths, along with the total number of patients, in the "Control" and the "Treated" group (with beta-blocker). Table

16 essentially provides individual-level data for a total of 20,290 patients. Taking Center 1 for example, there are 3 deaths among the 39 patients in the "Control" group. This corresponds to 39 observations with a treatment indicator variable being 0 ("Control") and a binary response variable being 1 for 3 subjects and 0 for the remaining 36. Following this argument, we can "reconstruct" individual observations (IPD data) of all the patients.

To conduct IPD analysis of the treatment effect, we model the probability of "Death" $P(Y = 1)$ in terms of the treatment indicator variable $X(= 0 \text{ or } 1)$. Specifically, we use a logistic regression model

$$P(Y_{ki} = 1|X_{ki}) = \frac{\exp(\alpha_k + \beta_k X_{ki})}{1 + \exp(\alpha_k + \beta_k X_{ki})} \quad (k = 1, \ldots, 22; i = 1, \ldots, n_k).$$

To allow possible heterogeneity among the centers, we assume that both the intercepts α_k's (i.e., the baseline effects) and the treatment effects β_k's are random and they follow a bivariate normal distribution

$$\begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \tau_\alpha^2 & \rho\tau_\alpha\tau_\beta \\ \rho\tau_\alpha\tau_\beta & \tau_\beta^2 \end{pmatrix} \right).$$

The maximum likelihood inference yields $\hat{\beta}_{IPD} = -0.247$ (i.e., the log odds ratio) with an estimated standard error of 0.057. This implies that the use of beta-blockade reduces the probability of death, and this effect is significant with a p-value of 1.65×10^{-5}.

For summary-statistics-based meta-analysis, we assume that only the center-specific treatment effect estimate $\hat{\beta}_k$ and its variance estimate $\widehat{\text{var}}(\hat{\beta}_k \mid \beta_k)$ are given from each center. These summary statistics are presented in Table 17. Plugging these statistic into (2.2), we carry out meta-analysis. The estimated treatment effect is $\hat{\beta}_{SS} = -0.250$ with a estimated standard error of 0.058, and we conclude that the treatment effect is significant with a p-value of 1.44×10^{-5}. The analysis results here are quite similar to those obtained from the IPD analysis. The similarity implies that our summary-statistics-based analysis virtually does not lose any efficiency, even in the situation where we are not given any information of the nuisance parameters α_k's. Our further examination shows that α_k's are, in fact, correlated with the treatment effects β_k's with a correlation estimate $\hat{\rho} = -0.45$. This result confirms our finding in Theorem 1(c), which says that meta-analysis using merely summary statistics of β_k's is fully efficient without the knowledge of the nuisance parameters α_k's even if the two random effects α_k's and β_k's are correlated.

Table 16: Beta-blocker data collected from 22 clinical centers

	Control		Treated			Control		Treated	
Center	Deaths	Total	Deaths	Total	Center	Deaths	Total	Deaths	Total
1	3	39	3	38	12	47	266	45	263
2	14	116	7	114	13	16	293	9	291
3	11	93	5	69	14	45	883	57	858
4	127	1520	102	1533	15	31	147	25	154
5	27	365	28	355	16	38	213	33	207
6	6	52	4	59	17	12	122	28	251
7	152	939	98	945	18	6	154	8	151
8	48	471	60	632	19	3	134	6	174
9	37	282	25	278	20	40	218	32	209
10	188	1921	138	1916	21	43	364	27	391
11	52	583	64	873	22	39	674	22	680

Table 17: Summary statistics used in meta-analysis of the Beta-blocker data

Center	$\hat{\beta}_k$	$\widehat{\text{var}}(\hat{\beta}_k \mid \beta_k)$	Center	$\hat{\beta}_k$	$\widehat{\text{var}}(\hat{\beta}_k \mid \beta_k)$
1	0.028	0.723	12	-0.039	0.053
2	-0.741	0.233	13	-0.593	0.181
3	-0.541	0.319	14	0.282	0.042
4	-0.246	0.019	15	-0.321	0.089
5	0.069	0.079	16	-0.135	0.068
6	-0.584	0.457	17	0.141	0.133
7	-0.512	0.019	18	0.322	0.305
8	-0.079	0.042	19	0.444	0.514
9	-0.424	0.075	20	-0.218	0.068
10	-0.335	0.014	21	-0.591	0.066
11	-0.213	0.038	22	-0.608	0.074

References

Freemantle, N., J. Cleland, P. Young, J. Mason, and J. Harrison (1999). Beta-blockade after myocardial infarction: systematic review and meta regression analysis. *BMJ 318*, 1730–1737.

Yusuf, S., R. Peto, J. Lewis, R. Collins, and P. Sleight (1985). Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progress in Cardiovascular Diseases. 27*, 335–371.