Optimal and maximin sample sizes for multicentre cost-effectiveness trials

Citation for published version (APA):

Manju, A., Candel, M. J. J. M., & Berger, M. P. F. (2015). Optimal and maximin sample sizes for multicentre cost-effectiveness trials. *Statistical Methods in Medical Research*, 24(5), 513-539. <https://doi.org/10.1177/0962280215569293>

Document status and date:

Published: 01/01/2015

DOI:

[10.1177/0962280215569293](https://doi.org/10.1177/0962280215569293)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Optimal and maximin sample sizes for multicentre cost-effectiveness trials

Md. Abu Manju, Math JJM Candel and
Martijn PF Berger

Statistical Methods in Medical Research
2015, Vol. 24(5) 513–539

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0962280215569293

smm.sagepub.com

Abstract

This paper deals with the optimal sample sizes for a multicentre trial in which the cost-effectiveness of two treatments in terms of net monetary benefit is studied. A bivariate random-effects model, with the treatment-by-centre interaction effect being random and the main effect of centres fixed or random, is assumed to describe both costs and effects. The optimal sample sizes concern the number of centres and the number of individuals per centre in each of the treatment conditions. These numbers maximize the efficiency or power for given research costs or minimize the research costs at a desired level of efficiency or power. Information on model parameters and sampling costs are required to calculate these optimal sample sizes. In case of limited information on relevant model parameters, sample size formulas are derived for so-called maximin sample sizes which guarantee a power level at the lowest study costs. Four different maximin sample sizes are derived based on the signs of the lower bounds of two model parameters, with one case being worst compared to others. We numerically evaluate the efficiency of the worst case instead of using others. Finally, an expression is derived for calculating optimal and maximin sample sizes that yield sufficient power to test the cost-effectiveness of two treatments.

Keywords

Cost-effectiveness, multicentre trials, maximin design, optimal design, power

I Introduction

The cost-effectiveness of interventions is often evaluated in multicentre randomized clinical trials, where persons within a centre are randomly assigned to one of two or more treatments. In such multicentre trials, individual level outcomes such as costs and effects are collected in a number of different centres. A prominent example is the EVALUATE trial,^{1,2} in which patients within each of the 25 centres were randomly assigned to one of two treatments, namely laparoscopic-assisted hysterectomy and standard hysterectomy. This study estimated the differences between treatment arms of mean effectiveness (incremental effectiveness) and mean costs (incremental costs), where

Department of Methodology and Statistics, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, The Netherlands

Corresponding author:

Md. Abu Manju, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

Email: abu.manju@maastrichtuniversity.nl

costs were expressed in terms of UK pounds (£) and effects in terms of quality-adjusted life years. Multicentre trials are very common in cost-effectiveness studies and will be denoted as multicentre cost-effectiveness trials in what follows.

In multicentre cost-effectiveness trials, individuals are nested within centres, for instance, patients are nested within hospitals. As a consequence, in such nested designs there may be variability in costs and effects between centres. A failure to address this variability may give misleading conclusions in the analysis of cost-effectiveness studies. In the literature on conventional multicentre efficacy trials, the two primary methods of accounting for such between-centre variability of outcomes in the analysis of data are fixed-effects and random-effects models.³⁻⁹ A major difference between them is the way in which they treat the main effect of centres and the treatment-by-centre interaction effect. In the fixed-effects model the main effect of centres and the treatment-by-centre interaction are regarded as being fixed, no different as other covariates such as age or gender. However, the random-effects model considers at least one of these two effects a random variable that follows a distribution (usually a normal distribution).⁴ Treating the main effect of centres as random, as illustrated by Kahan and Morris,⁷ can improve the precision of the treatment estimates compared to when the main effect of centres is treated as fixed. However, in case of perfect balance (equal numbers of patients in a treatment group for each centre), this precision of the treatment estimates is the same regardless of whether the main effect of centres is treated as random or fixed.⁹

A random-effects model in which the treatment effect is assumed to vary randomly from centre to centre by treating the treatment-by-centre interaction as random has, for example been studied by Fedorov and Jones.³ Such a random-effects model allows to generalize the inferences beyond the centres used in the trial and assumes that the centres are a random sample from a larger population of centres.^{4,5} Of course, in practice the centres in trials are often chosen in some logical way and in many cases are not a random sample from a normal population.⁵ However, if the treatment-by-centre interaction is treated as fixed, the inferences from such a model strictly are only applicable to the set of centres used in the trial.^{4,5} Therefore, these inferences do not tell us what will happen if the wider issue of the probable effect of a treatment is being considered. The study by Dragalin et al.⁶ is an example of a fixed-effects approach in a multicentre trial.

The issue of whether to fit a fixed- or random-effects model has been a controversial and much disputed subject within the field of multicentre trials. This has resulted in much debate and discussion in the literature.^{4,5} However, a number of researchers have reported that the random-effects model performed as well as or better than the fixed-effects model to address the variation observed in data collected in multicentre trials.^{3,5,7} For example, a random-effects model offers advantages over a fixed-effects model in terms of its ability to account for centre level covariates and to calculate centre-specific estimates of a treatment's effectiveness through estimating shrunken residuals. In several studies, in which data from multicentre cost-effectiveness trials were analysed, the random-effects model was used, treating the treatment-by-centre interaction as random and the main effect of centres either as fixed or random.^{2,10-17} In this paper, we will therefore start from a random-effects model treating the treatment-by-centre interaction as random, which in the literature also is known as a mixed-effects model.¹⁸ Since we start from a balanced design, in Section 3 we will see that, independent of whether the main effect of centres is fixed or random, the same expression for the treatment effect estimator and its standard error result.

The planning of multicentre cost-effectiveness trials involves several design choices, including sample size per treatment arm in each centre and the number of centres. There have been many developments of statistical methods for determining the sample sizes of multicentre cost-effectiveness trials.¹⁹⁻²³ However, these methods ignore the between centre variability of cost-effectiveness data

which is an important characteristic of multicentre cost-effectiveness trials. Ignoring this variability of outcomes results in an incorrect expression for the variance of the treatment effect estimator and consequently results in inadequate sample sizes and statistical power.^{24,25}

This paper is about efficient sample size calculation in multicentre cost-effectiveness trials, which recognizes the nesting feature of cost-effectiveness data and accounts for variability among the responses of individuals between centres. In a well-planned trial the sample sizes should be considered that minimize the total costs for sampling and measuring centres and persons within the centres while the trial meets required statistical power. Alternatively, the sample sizes are obtained by maximizing the efficiency of the treatment effect estimator, and thus results in maximum power for a Wald test of the treatment effect, given a fixed research budget. These sample sizes are referred to as optimal sample sizes, also denoted as the optimal design.

Guidelines for calculating optimal sample sizes in multicentre trials which account for the nested structure of the data are available, but these optimal sample sizes are only based on the effectiveness of interventions.^{3,25–27} As an example, Fedorov and Jones,³ using the random-effects model, determined the optimal number of centres and the optimal number of patients per centre assuming the costs per sampling unit are equal between treatment arms. The current paper presents optimal sample sizes for multicentre cost-effectiveness trials allowing the costs per sampling unit to vary between treatments. To accommodate for both costs and effects of interventions in deriving optimal sample sizes, net monetary benefit (*NMB*)^{2,10} is taken as the outcome variable. Starting from the structure of a bivariate random-effects model, these costs and effects are assumed to follow a bivariate normal distribution.^{2,10} The model can thus be considered a bivariate version of the random-effects model 5.1 or 5.2 as described by Senn⁴ (see p. 539), where the treatment-by-centre interaction is random and the main effect of centres is either fixed or random.

The construction of the optimal design requires the advance knowledge of model parameters, which may be unknown at the planning stage of a trial. This is called the local optimality problem.^{28,29} Three approaches have been proposed in the literature to deal with the local optimality problem, namely the sequential approach, the Bayesian approach and the maximin approach. The sequential approach improves the design based on the improved parameter estimates and it usually requires a long time to obtain an efficient design.³⁰ The Bayesian approach results in a design which maximizes the power for testing the treatment effect by integrating over prior distributions of the parameters.³¹ This approach assumes prior distributions for the parameters and performs well on average. Finally, the maximin approach maximizes performance under the worst-case scenario in terms of the unknown parameters.²⁸ This approach only needs to specify plausible ranges for the unknown parameters and guarantees a certain level of efficiency for a prespecified range of possible parameter values. Thus to accommodate the local optimality problem, the maximin approach has the best possible worst-case performance compared with the Bayesian and the sequential approaches. To solve this local optimality problem, the current paper presents maximin sample sizes for multicentre cost-effectiveness trials.

This paper is structured as follows. Section 2 provides a brief description of the modelling framework for multicentre cost-effectiveness trials. In Section 3, we derive the optimal sample sizes for multicentre cost-effectiveness trials, which maximize the efficiency of the estimator of the average incremental *NMB* for comparing two treatments, for a given cost function specifying the research costs. In Section 4, maximin sample sizes are developed to accommodate the case when information on the model parameters is rather vague. A comparison among the maximin sample sizes, which are derived based on the signs of the parameter boundaries, is done in Section 5. Next, we propose a formula which describes the connection between the sample sizes and power of the test

on the cost-effectiveness of a treatment. In Section 6, we apply our proposed sample sizes to a case study and discuss the findings. The paper concludes with a summary of our results, points out limitations of the present method and gives suggestions for further research.

2 Model specification

We present a general modelling framework for balanced multicentre cost-effectiveness trials. To describe the models, let E_{ij} and C_{ij} represent the effects and costs for individual i in centre j ($j = 1, 2, \dots, k$), where $i = 1, 2, \dots, m$ for an individual in the treatment condition, and $i = m + 1, m + 2, \dots, m + n$ for an individual in the control condition, in centre j . We start from the following bi-variate linear mixed model for multicentre cost-effectiveness trials¹⁰

$$E_{ij} = \beta_{0j}^e + \beta_1^e x_{ij} + u_{1j}^e x_{ij} + \varepsilon_{ij}^e \quad \text{and} \quad C_{ij} = \beta_{0j}^c + \beta_1^c x_{ij} + u_{1j}^c x_{ij} + \varepsilon_{ij}^c$$

$$\begin{pmatrix} u_{1j}^e \\ u_{1j}^c \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{e1}^2 & \rho_{u1}\sigma_{e1}\sigma_{c1} \\ & \sigma_{c1}^2 \end{pmatrix}\right), \quad \text{and} \quad \begin{pmatrix} \varepsilon_{ij}^e \\ \varepsilon_{ij}^c \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon^e}^2 & \rho_{\varepsilon}\sigma_{\varepsilon^e}\sigma_{\varepsilon^c} \\ & \sigma_{\varepsilon^c}^2 \end{pmatrix}\right) \quad (1)$$

where x_{ij} takes the values a or b (a and b are real numbers and $a \neq b$) according to whether the i th individual in centre j is in the control or treatment condition, that is $x_{ij} = a$ for control and $x_{ij} = b$ for treatment. The common coding schemes (a, b) in the literature are $0/1$, $-1/1$ and $-0.5/0.5$ and we want to derive general expressions for sample size formulas which hold for any coding scheme. The intercepts β_{0j}^e and β_{0j}^c for centre j are the main effect of centres for costs and effects, respectively, and may be fixed or random. In terms of equation (1), the incremental costs and effects for centre j are $(b - a)(\beta_1^c + u_{1j}^c)$ and $(b - a)(\beta_1^e + u_{1j}^e)$, respectively. The random slopes u_{1j}^c and u_{1j}^e represent the treatment-by-centre interaction for costs and effects, respectively. These interaction effects are assumed to come from a bivariate normal distribution with means of zero, variances σ_{c1}^2 and σ_{e1}^2 and correlation ρ_{u1} ($-1 \leq \rho_{u1} \leq 1$). The bivariate model in equation (1) also allows for individual level variances $\sigma_{\varepsilon^e}^2$ and $\sigma_{\varepsilon^c}^2$ and a correlation between costs and effects at the individual level ρ_{ε} ($-1 \leq \rho_{\varepsilon} \leq 1$). Let denote $\sigma_{e\bar{c}}^2 = \sigma_{e1}^2 + (b - a)^{-2}\sigma_{\varepsilon^e}^2$, and $\sigma_c^2 = \sigma_{c1}^2 + (b - a)^{-2}\sigma_{\varepsilon^c}^2$ for further simplicity, and define, $0 \leq \theta_e = \frac{\sigma_{e1}^2}{\sigma_{e\bar{c}}^2} \leq 1$ and $0 \leq \theta_c = \frac{\sigma_{c1}^2}{\sigma_c^2} \leq 1$ where θ_e (or, θ_c) represents the proportion of the variability of incremental effects (or, costs) from treatment, $(b - a)u_{1j}^e$, (or, $(b - a)u_{1j}^c$) compared to the variability of effects, $(b - a)u_{1j}^e + \varepsilon_{ij}^e$ (or, costs, $(b - a)u_{1j}^c + \varepsilon_{ij}^c$).

The models in equation (1) are used to estimate average incremental effects and costs when data on effects and costs for each individual are available from a multicentre trial and these estimates are on their original scales. The concept of *NMB*, where both the effects and costs of interventions are scaled onto the same monetary scale on the basis of a threshold value λ , offers a convenient way of modelling the data from multicentre costs-effectiveness trials.³² The value λ ($0 \leq \lambda < \infty$) is the societal willingness to pay for a unit of health gain or the maximum value associated with an additional unit of effectiveness.³² The *NMB* for each individual i in the j th centre can be calculated as²

$$NMB_{ij} = \lambda E_{ij} - C_{ij} \quad (2)$$

Equation (2) can be elaborated as

$$NMB_{ij} = \lambda(\beta_{0j}^e + \beta_1^e x_{ij} + u_{1j}^e x_{ij} + \varepsilon_{ij}^e) - (\beta_{0j}^c + \beta_1^c x_{ij} + u_{1j}^c x_{ij} + \varepsilon_{ij}^c) = \beta_{0j} + \beta_1 x_{ij} + u_{1j} x_{ij} + \varepsilon_{ij} \quad (3)$$

where $\beta_{0j} (= \lambda\beta_{0j}^e - \beta_{0j}^c)$ is the intercept term for centre j and x_{ij} is a variable for the treatment condition with coefficient $\beta_1 (= \lambda\beta_1^e - \beta_1^c)$. The average incremental *NMB* (*INMB*) from treatment is $(b - a)\beta_1$. The intervention should replace the control treatment if $(b - a)\beta_1 > 0$, that is a positive *INMB* means that the intervention is cost-effective. It follows that $u_{1j} (= \lambda u_{1j}^e - u_{1j}^c) \sim N(0, \sigma_1^2)$ and $\varepsilon_{ij} (= \lambda \varepsilon_{ij}^e - \varepsilon_{ij}^c) \sim N(0, \sigma^2)$, where $\sigma_1^2 = (\phi\theta_e + \theta_c - 2\rho_{u1}\sqrt{\theta_e\theta_c\phi})\sigma_c^2$ and $\sigma^2 = (b - a)^2\{\phi(1 - \theta_e) + (1 - \theta_c) - 2\rho_\varepsilon\sqrt{(1 - \theta_e)(1 - \theta_c)\phi}\}\sigma_c^2$. We define $\phi = \frac{\lambda^2\{(b-a)^2\sigma_{e1}^2 + \sigma_{\varepsilon e}^2\}}{\{(b-a)^2\sigma_{c1}^2 + \sigma_{\varepsilon c}^2\}}$, where the numerator and denominator of this ratio are the variance of random effects, $(b - a)u_{1j}^e + \varepsilon_{ij}^e$, translated into costs, $\lambda\{(b - a)u_{1j}^e + \varepsilon_{ij}^e\}$, and the variance of the random costs, $(b - a)u_{1j}^c + \varepsilon_{ij}^c$, respectively. This ϕ ($0 \leq \phi < \infty$) can be interpreted as the variance of the effects translated into costs $\lambda\{(b - a)u_{1j}^e + \varepsilon_{ij}^e\}$ of obtaining one additional unit of variance in costs $\{(b - a)u_{1j}^c + \varepsilon_{ij}^c\}$. As $\lambda \rightarrow \infty$ (or, $\lambda = 0$), equation (3) reduces to the first part of equation (1), that is the regression model of effects on treatment (or, the second part of equation (1), that is the regression model of costs on treatment). Note that each of the model parameters θ_e , θ_c , ρ_{u1} , ρ_ε and ϕ is invariant under a linear transformation of the variable x_{ij} .

3 Optimal sample sizes

It is assumed in deriving the optimal sample sizes that the variances of the random components in equation (3) are positive, that is $\sigma_1^2 = (\phi\theta_e + \theta_c - 2\rho_{u1}\sqrt{\theta_e\theta_c\phi})\sigma_c^2 > 0$ and $\sigma^2 = (b - a)^2\{\phi(1 - \theta_e) + (1 - \theta_c) - 2\rho_\varepsilon\sqrt{(1 - \theta_e)(1 - \theta_c)\phi}\}\sigma_c^2 > 0$, otherwise our derivations do not hold. So the parameter values for θ_e , θ_c , ρ_{u1} , ρ_ε , ϕ and σ_c^2 have to be chosen such that $\sigma_1^2 > 0$ and $\sigma^2 > 0$. For the linear mixed model in equation (3), the maximum likelihood estimator of the treatment effect β_1 is given by¹⁸

$$\hat{\beta}_1 = \frac{1}{k(b - a)} \left(\frac{1}{m} \sum_{j=1}^k \sum_{i=1}^m NMB_{ij} - \frac{1}{n} \sum_{j=1}^k \sum_{i=m+1}^{m+n} NMB_{ij} \right) = \frac{\hat{\mu}_1 - \hat{\mu}_2}{b - a} \quad (4)$$

in which $\hat{\mu}_1$ and $\hat{\mu}_2$ are the mean responses in the treatment and control condition, respectively. The variance of the treatment effect estimator in equation (4) can be derived based on maximum likelihood estimation as covered by Verbeke and Molenberghs¹⁸

$$Var(\hat{\beta}_1) = \frac{1}{k} \left\{ \sigma_1^2 + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^2}{(b - a)^2} \right\} \quad (5)$$

Thus, the variance of the treatment effect estimator ($\hat{\beta}_1$) depends on the model parameters θ_e , θ_c , ρ_{u1} , ρ_ε , ϕ and σ_c^2 ; the number of individuals m and n in the treatment and control condition, respectively, per centre; and the number of centres k . Note that the term $(b - a)^{-2}\sigma^2$, appears in the expression for $Var(\hat{\beta}_1)$ in equation (5), which, given the definition of σ^2 (see Section 2), is independent of the treatment coding (a , b). This means that different coding schemes will lead to the same expression for $Var(\hat{\beta}_1)$ in equation (5).

From equation (5) it is clear that increasing either the centre sizes m and n or the number of centres k decreases $Var(\hat{\beta}_1)$, and thus maximizes the statistical power and the precision. Given certain costs attached to the number of subjects and the number of centres, the issue is what would be the best choice for m , n and k . We can address this problem by adopting optimal design theory. The optimal sample sizes minimize the $Var(\hat{\beta}_1)$ in equation (5), and thus maximize the statistical power for the test of $\hat{\beta}_1$ in equation (4), for a given constraint on sampling and measuring costs attached to centres and individuals. To find the optimal sample sizes we start from the following budget constraint²⁷

$$C = k(c_0 + c_1m + c_2n) \quad (6)$$

where $c_0 > 0$ is the cost per included centre, $c_1 > 0$ and $c_2 > 0$ are the costs per included individual in the treatment and control condition, respectively, and most likely we have $c_0 \geq c_1$ and $c_0 \geq c_2$.^{26,33} Finally, C stands for the total budget.

The following optimal sample sizes are obtained by minimizing the variance of the treatment effect estimator in equation (5) for a fixed total budget C in equation (6)

$$m_{opt} = \sqrt{\frac{c_0\sigma^2}{c_1(b-a)^2\sigma_1^2}}, \quad n_{opt} = \sqrt{\frac{c_0\sigma^2}{c_2(b-a)^2\sigma_1^2}} \quad \text{and} \quad k_{opt} = \frac{C}{\sqrt{\frac{c_0\sigma^2}{(b-a)^2\sigma_1^2}}(\sqrt{c_1} + \sqrt{c_2}) + c_0} \quad (7)$$

The optimal centre sizes m_{opt} and n_{opt} increase as the centre-to-individual cost ratios c_0/c_1 and c_0/c_2 increase, respectively. The budget C affects the optimal number of centres, k_{opt} , but not the optimal centre sizes. As discussed for $Var(\hat{\beta}_1)$ below in equation (5), though equation (7) shows that the optimal sample sizes are a function of the treatment coding (a , b), it is the case that $(b-a)^{-2}\sigma^2$ does not depend on the coding, and therefore also not the optimal sample sizes. Also note that Liu²⁷ derived the optimal sample sizes for a linear mixed model for effects only which are the special case of equation (7), resulting by letting $\lambda \rightarrow \infty$ (and thus $\phi \rightarrow \infty$) and $|b-a| = 1$. Substituting the optimal sample sizes in equation (7) into equation (5) yields the following optimal variance for a vector of parameters $\Pi = (\theta_e, \theta_c, \rho_{u1}, \rho_e, \phi, \sigma_c^2)$

$$Var(\hat{\beta}_1|\Pi, opt) = \frac{\{\sigma(\sqrt{c_1} + \sqrt{c_2}) + \sqrt{c_0}(b-a)\sigma_1\}^2}{(b-a)^2C} \quad (8)$$

Equation (8) shows that the optimal variance decreases as the total budget, C , goes up and increases as c_1 , c_2 and c_0 increase.

4 Maximin sample sizes

Equation (7) shows that the optimal sample sizes are functions of the model parameters θ_e , θ_c , ρ_{u1} , ρ_e , ϕ and σ_c^2 . So information on the model parameters is necessary to obtain the optimal design. Unfortunately, the optimal design for one set of parameter values can be inefficient for another set of parameter values. The optimal design is therefore said to be locally optimal. In the present study, we consider the maximin approach to solve this local

optimality problem. The maximin approach takes the following two steps to obtain maximin sample sizes:

(a) minimize the following efficiency across the plausible ranges for the model parameters

$$Eff = \left\{ Var(\hat{\beta}_1) \right\}^{-1} = k \left\{ \sigma_1^2 + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^2}{(b-a)^2} \right\}^{-1} \quad (9)$$

where σ_1^2 and σ^2 are defined in Section 2, and

(b) choose that sample size (m, n, k) which maximizes this minimum efficiency given a fixed budget.

This is called the maximin design (MMD).

The efficiency in equation (9) depends on the model parameters θ_e , θ_c , ρ_{u1} , ρ_ε , ϕ and σ_c^2 . To obtain the minimum efficiency, we assume some plausible ranges for the parameters, $\rho_{u1} \in [\rho_{u1}^L, \rho_{u1}^U]$ with $-1 \leq \rho_{u1}^L \leq \rho_{u1}^U \leq 1$, $\rho_\varepsilon \in [\rho_\varepsilon^L, \rho_\varepsilon^U]$ with $-1 \leq \rho_\varepsilon^L \leq \rho_\varepsilon^U \leq 1$, $\theta_e \in [\theta_e^L, \theta_e^U]$ with $0 \leq \theta_e^L \leq \theta_e^U \leq 1$, $\theta_c \in [\theta_c^L, \theta_c^U]$ with $0 \leq \theta_c^L \leq \theta_c^U \leq 1$, $\phi \in [\phi^L, \phi^U]$, with $0 \leq \phi^L \leq \phi^U < \infty$ and $\sigma_c^2 \in [\sigma_c^{2L}, \sigma_c^{2U}]$, with $0 \leq \sigma_c^{2L} \leq \sigma_c^{2U} < \infty$. We can derive different maximin sample sizes depending on signs of the lower bounds for ρ_{u1} and ρ_ε . Figure 1 distinguishes the different cases under which maximin sample sizes are derived (see Appendix 1 for details).

Figure 1. Overview of the cases for which maximin designs are derived, classified according to the signs of lower bounds for ρ_{u1} and ρ_ε respectively. For all cases the maximin designs are different, except for Case 4. The maximin design for Case 4 reduces to the maximin design for Case 1b.

Case 1a ($-1 \leq \rho_{u1}^L < \rho_\varepsilon^L < 0$): The MMD (m_{MMD} , n_{MMD} , k_{MMD}) and the variance for the MMD , the maximin variance (MMV), become

$$m_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_1(b-a)^2 \sigma_1^{*2}}}, \quad n_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_2(b-a)^2 \sigma_1^{*2}}}, \quad k_{MMD} = \frac{C}{\sqrt{\frac{c_0 \sigma^{*2}}{(b-a)^2 \sigma_1^{*2}}(\sqrt{c_1} + \sqrt{c_2}) + c_0}}, \quad \text{and}$$

$$Var(\hat{\beta}_1 | \Pi^*, MMD) = MMV = \frac{\{\sigma^*(\sqrt{c_1} + \sqrt{c_2}) + \sqrt{c_0}(b-a)\sigma_1^*\}^2}{(b-a)^2 C} \quad (10)$$

where $\sigma_1^{*2} = \theta(\phi^U + 1 - 2\rho_{u1}^L \sqrt{\phi^U})\sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2(1-\theta)(\phi^U + 1 - 2\rho_\varepsilon^L \sqrt{\phi^U})\sigma_c^{2U}$, $\theta = \max\{\theta_e^U, \theta_c^U\}$ and $\Pi^* = (\theta, \theta, \rho_{u1}^L, \rho_\varepsilon^L, \phi^U, \sigma_c^{2U})$.

Case 1b ($-1 \leq \rho_\varepsilon^L \leq \rho_{u1}^L < 0$):

The following MMD and its MMV result

$$m_{MMD} = \sqrt{\frac{c_0(1-\theta)}{c_1\theta}}, \quad n_{MMD} = \sqrt{\frac{c_0(1-\theta)}{c_2\theta}}, \quad k_{MMD} = \frac{C}{\sqrt{\frac{c_0(1-\theta)}{\theta}(\sqrt{c_1} + \sqrt{c_2}) + c_0}}, \quad \text{and}$$

$$Var(\hat{\beta}_1 | \Pi^*, MMD) = MMV = \frac{\{\sigma^*(\sqrt{c_1} + \sqrt{c_2}) + \sqrt{c_0}(b-a)\sigma_1^*\}^2}{(b-a)^2 C} \quad (11)$$

where $\sigma_1^{*2} = \theta(\phi^U + 1 - 2\rho\sqrt{\phi^U})\sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2(1-\theta)\{\phi^U + 1 - 2\rho\sqrt{\phi^U}\}\sigma_c^{2U}$, $\rho = \min\{\rho_{u1}^L, \rho_\varepsilon^L\}$ and $\Pi^* = (\theta, \theta, \rho, \rho, \phi^U, \sigma_c^{2U})$.

Case 2 ($\rho_{u1}^L \geq 0$ and $\rho_\varepsilon^L \geq 0$): The maximin design and the MMV become

$$m_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_1(b-a)^2 \sigma_1^{*2}}}, \quad n_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_2(b-a)^2 \sigma_1^{*2}}}, \quad k_{MMD} = \frac{C}{\sqrt{\frac{c_0 \sigma^{*2}}{(b-a)^2 \sigma_1^{*2}}(\sqrt{c_1} + \sqrt{c_2}) + c_0}}$$

and $Var(\hat{\beta}_1 | \Pi^*, MMD) = MMV = \frac{\{\sigma^*(\sqrt{c_1} + \sqrt{c_2}) + \sqrt{c_0}(b-a)\sigma_1^*\}^2}{(b-a)^2 C} \quad (12)$

where $\sigma_1^{*2} = (\phi^U \theta_e^U + \theta_c^U)\sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2\{\phi^U(1-\theta_e^U) + (1-\theta_c^U)\}\sigma_c^{2U}$ and $\Pi^* = (\theta_e^U, \theta_c^U, 0, 0, \phi^U, \sigma_c^{2U})$.

Case 3 ($\rho_{u1}^L < 0$ and $\rho_\varepsilon^L \geq 0$):
The MMD becomes

$$m_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_1(b-a)^2 \sigma_1^{*2}}}, \quad n_{MMD} = \sqrt{\frac{c_0 \sigma^{*2}}{c_2(b-a)^2 \sigma_1^{*2}}}, \quad k_{MMD} = \frac{C}{\sqrt{\frac{c_0 \sigma^{*2}}{(b-a)^2 \sigma_1^{*2}}(\sqrt{c_1} + \sqrt{c_2}) + c_0}}$$

and the corresponding MMV is given by

$$Var(\hat{\beta}_1|\Pi^*, MMD) = MMV = \frac{\{\sigma^*(\sqrt{c_1} + \sqrt{c_2}) + \sqrt{c_0}(b-a)\sigma_1^*\}^2}{(b-a)^2C} \quad (13)$$

where $\sigma_1^{*2} = (\phi^U\theta_e^U + \theta_c^U - 2\rho_{u1}^L\sqrt{\phi^U\theta_e^U\theta_c^U})\sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2\{\phi^U(1-\theta_e^U) + (1-\theta_c^U)\}\sigma_c^{2U}$ and $\Pi^* = (\theta_e^U, \theta_c^U, \rho_{u1}^L, 0, \phi^U, \sigma_c^{2U})$.

Case 4 ($\rho_{u1}^L \geq 0$ and $\rho_\varepsilon^L < 0$):

In this case our derivation does not provide any simple solution for the MMV if we start with $\rho_{u1}^L = 0$. Starting with a positive lower bound for ρ_{u1} , that is $\rho_{u1}^L > 0$ does also not give a solution for the maximin design. In that case, assuming a small negative value for ρ_{u1}^L , that is $-1 \leq \rho_\varepsilon^L < \rho_{u1}^L < 0$ might be the solution to get a maximin design, which corresponds to Case 1b (see Equation (11)).

The MMV increases with increasing the costs for each individual (c_1, c_2) and the centre-specific costs (c_0), but decreases with the total budget, C . Note that the researcher should take the upper bound for a plausible range of σ_c^2 in calculating the MMV in equations (10) to (13). However, as we will see in Section 6, when calculating k_{MMD} such that the MMV is small enough to obtain a desired power level in testing, use can be made of an effect size. The smallest effect size which is considered still clinically important could be chosen. In that case an upper bound for σ_c^2 no longer needs to be specified (see Section 6 for details). Since the maximin sample sizes in equations (10) to (13) are the optimal sample sizes in equation (7) for parameter boundaries, the maximin sample sizes also do not depend on the treatment coding. Note that the maximin sample sizes in equations (10) to (13) are restricted to parameter boundaries such that $\sigma_1^{*2} > 0$ and $\sigma^{*2} > 0$.

5 Relative efficiencies of the maximin sample sizes

We have found four different maximin designs based on the signs of the lower bounds for the correlations ρ_{u1} and ρ_ε . The $MMDs$ under Case 1 (a and b) are based on the worst cases compared to the $MMDs$ in Case 2 and Case 3. Also Case 1b is based on a case worse than Case 1a. The researcher should have specific information on the lower bounds of ρ_{u1} and ρ_ε to calculate the appropriate MMD . On the other hand, if there is vague knowledge on ρ_{u1} and ρ_ε , one can always start from Case 1b, as, in line with the maximin approach, we then optimize the worst-case scenario. The question is how much efficiency we lose if we start from the worst case (Case 1b) but Case 1a, Case 2 and Case 3 are the true cases. To measure this efficiency loss we define the following relative efficiency

$$RE_{Case\ 1b|Case\ h} = \frac{Var(\hat{\beta}_1|\Pi_{Case\ h}^*, MMD_{Case\ h})}{Var(\hat{\beta}_1|\Pi_{Case\ h}^*, MMD_{Case\ 1b})} \quad (14)$$

where $h = 1a, 2, 3$ and for each value of h , and $\Pi_{Case\ h}^*$ and $MMD_{Case\ h}$ are defined as the parameter vector and the maximin design, respectively, for Case h . This ratio measures the efficiency loss when the MMD under Case 1b is assumed, while instead the MMD by Case h should have been used. The $RE_{Case\ 1b|Case\ h}$ in equation (14) depends on the parameter boundaries under which the $MMDs$ in

Case *h* and Case 1*b* are derived and also on the cost ratios c_0/c_1 and c_0/c_2 , but the total budget, C , and the parameter boundary σ_c^{2U} cancel out (see supplementary materials for details [available at <http://smm.sagepub.com>]). We can examine the following three relative efficiencies of the *MMD* by Case 1*b* versus the *MMD* by Case *h* with $h = 1a, 2, 3$:

Case 1*b* versus Case 1*a*

$RE_{Case\ 1b|Case\ 1a}$ refers to the relative efficiency of the worst Case 1*b* compared to Case 1*a*. This relative efficiency depends on the parameter boundaries θ , ρ_ε^L , ρ , ϕ^U and the cost ratios c_0/c_1 and c_0/c_2 . We investigated $RE_{Case\ 1b|Case\ 1a}$ numerically for a number of different realistic parameter boundaries and costs ratios. Figures 2 and 3 present $RE_{Case\ 1b|Case\ 1a}$ as a function of ϕ^U , which ranges from 0.1 to 10, for three different values of costs ratios and with eight different values of ρ_ε^L , for $\theta = 0.2$ and 0.8, respectively. Evaluating $RE_{Case\ 1b|Case\ 1a}$ we always consider $\rho = \min\{\rho_{u1}^L, \rho_s^L\} = -1.0$, because maximum efficiency loss will occur in this case. All the figures show that the $RE_{Case\ 1b|Case\ 1a}$ is never below 0.97. For other values of θ , $RE_{Case\ 1b|Case\ 1a}$ shows the same pattern as in Figures 2 and 3. It can also be seen from Figures 2 and 3 that the relative efficiency decreases as ρ_ε^L goes up. Furthermore, the results of different numerical evaluations for many different values of θ show that the $RE_{Case\ 1b|Case\ 1a}$ increases as the cost ratios increase provided $\theta \geq 0.5$ (as illustrated by Figure 3). On the other hand, if $\theta < 0.5$ the $RE_{Case\ 1b|Case\ 1a}$ does not consistently increase or decrease with the costs ratios (as illustrated by Figure 2).

Case 1*b* versus Case 2

The relative efficiency $RE_{Case\ 1b|Case\ 2}$ is a function of θ_e^U , θ_c^U , θ , ϕ^U and the costs ratios c_0/c_1 and c_0/c_2 . Since $\theta = \max\{\theta_e^U, \theta_c^U\}$, there are two possibilities either $\theta = \theta_e^U$ or $\theta = \theta_c^U$. This means that if $\theta = \theta_e^U$ (or, $\theta = \theta_c^U$), then $\theta_e^U > \theta_c^U$ (or, $\theta_c^U > \theta_e^U$). We can analytically prove that $RE_{Case\ 1b|Case\ 2}$ has the following properties (see supplementary materials [available at <http://smm.sagepub.com>]):

- (i) If $\theta = \theta_e^U > \theta_c^U$ the relative efficiency $RE_{Case\ 1b|Case\ 2}$ decreases as ϕ^U decreases, whereas if $\theta = \theta_c^U > \theta_e^U$ then the $RE_{Case\ 1b|Case\ 2}$ decreases as ϕ^U increases. The $RE_{Case\ 1b|Case\ 2}$ becomes 1.0 when $\theta = \theta_e^U = \theta_c^U$.
- (ii) The relative efficiency $RE_{Case\ 1b|Case\ 2}$ is independent of $\rho = \min\{\rho_{u1}^L, \rho_\varepsilon^L\}$ defined in Case 1*b*.

We examined $RE_{Case\ 1b|Case\ 2}$ numerically and for this we use some plausible boundaries of the parameters and cost ratios. In Figures 4 to 7, the relative efficiencies, $RE_{Case\ 1b|Case\ 2}$ over $\phi^U \in [0.1, 10]$ are presented for different values of θ_e^U , θ_c^U , θ , ϕ^U , c_0/c_1 and c_0/c_2 . Figures 4 and 5 represent the relative efficiency as a function of ϕ^U where it is assumed $\theta = \theta_e^U > \theta_c^U$ and in Figures 6 and 7, we assume $\theta = \theta_c^U > \theta_e^U$ in calculating the relative efficiency. In Figures 4 and 5, the parameter boundary θ_c^U varies across plots and for specified values of $\theta = \theta_e^U$ and θ_c^U , each plot of Figures 4 and 5 shows the relative efficiencies over ϕ^U for three different cost ratios indicated by different line types. For example, the upper left plot of Figure 4 corresponds to $\theta = \theta_e^U = 0.2$, $\theta_c^U = 0.175$. Furthermore, Figures 6 and 7 show similar relative efficiency plots for varying $\phi^U \in [0.1, 10]$ and for three different cost ratios. In these figures, the parameter boundary θ_e^U varies across the plots, while the boundary $\theta = \theta_c^U$ varies across the figures.

Several patterns emerge from these figures. First, as proven (see (i)), if $\theta = \theta_e^U > \theta_c^U$ the relative efficiency $RE_{Case\ 1b|Case\ 2}$ decreases as ϕ^U decreases (see Figures 4 and 5), whereas if $\theta = \theta_c^U > \theta_e^U$

Figure 2. Efficiency of the MMD under Case 1b compared to the MMD by Case 1a as a function of the parameter ϕ^U for three different sets of cost ratios and eight different values of ρ_ε^L assuming θ and ρ_{u1}^L are fixed at 0.2 and -1.0, respectively.

$RE_{Case\ 1b|Case\ 2}$ decreases as ϕ^U increases (see Figures 6 and 7). Second, from all figures, it is seen that larger absolute differences between θ_e^U and θ_c^U result in larger efficiency losses. Third, the $RE_{Case\ 1b|Case\ 2}$ decreases as the cost ratios increase in the case $\theta < 0.5$ (see Figures 4 and 6). As the costs of selecting a new centre gets higher as compared to the costs of an additional individual per centre, the MMD under Case 1b will lose efficiency provided $\theta = \max\{\theta_e^U, \theta_c^U\} < 0.5$. However, no simple pattern of $RE_{Case\ 1b|Case\ 2}$ seems to emerge when changing the cost ratios if $\theta \geq 0.5$ (see Figures 5 and 7).

Figure 3. Efficiency of the *MMD* under *Case 1b* compared to the *MMD* by *Case 1a* as a function of the parameter ϕ^U for three different sets of cost ratios and eight different values of ρ_ε^L assuming θ and ρ_{u1}^L are fixed at 0.8 and -1.0, respectively.

Case 1b versus Case 3

Finally, we consider the efficiency of the *MMD* in *Case 1b* compared to the *MMD* in *Case 3*, $RE_{Case\ 1b|Case\ 3}$. The parameter boundaries θ_e^U , θ_c^U , θ , ρ_{u1}^L , ρ , ϕ^U and the costs ratios c_0/c_1 and c_0/c_2 play a role in the comparison of these two *MMDs*. We plotted $RE_{Case\ 1b|Case\ 3}$ as a function of $\phi^U \in [0.1, 10]$ for some realistic parameter boundaries, which are shown in Figures 8 to 11. In Figures 8 and 9, we assume $\theta = \theta_e^U > \theta_c^U$ and in this case $RE_{Case\ 1b|Case\ 3}$ increases, with increasing ϕ^U , while in Figures 10 and 11, where it is assumed that $\theta = \theta_c^U > \theta_e^U$, the $RE_{Case\ 1b|Case\ 3}$ decreases with

Figure 4. Efficiency of the MMD under Case 1b compared to the MMD by Case 2 as a function of the parameter ϕ^U for three different set of cost ratios assuming $\theta = \theta_e^U$ is fixed at 0.20 and θ_c^U varies.

increasing ϕ^U . From Figures 8 to 11, it can also be seen that the trends of the $RE_{Case\ 1b|Case\ 3}$, as the cost ratios increase, are similar with the previously discussed relative efficiency $RE_{Case\ 1b|Case\ 2}$. All the figures are evaluated by considering $\rho = \min\{\rho_{u1}^L, \rho_\varepsilon^L\} = -1$ and $\rho_{u1}^L = -0.001$ because the most negative values for ρ and the least negative values for ρ_{u1}^L result in maximum efficiency loss.

The purpose of studying the relative efficiency, $RE_{Case\ 1b|Case\ h}$ for $h = 1a, 2, 3$ is to answer the question whether the researcher must have to know to which of the regions of Figure 1 the lower bounds of ρ_{u1} and ρ_ε belong to apply a MMD or whether one can start, without loss of efficiency, with the worst case MMD (Case 1b). Since we do not have much more empirical information on the

Figure 5. Efficiency of the MMD under Case 1b compared to the MMD by Case 2 as a function of the parameter ϕ^U for three different sets of cost ratios assuming $\theta = \theta_e^U$ is fixed at 0.8 and θ_c^U varies.

parameters for multicentre cost-effectiveness trials, we have considered some hypothetical boundaries for the parameters in our numerical study evaluating $RE_{Case\ 1b|Case\ h}$ for $h = 1a, 2, 3$. The numerical evaluation shows that the relative efficiency $RE_{Case\ 1b|Case\ 1a}$ is greater than or equal to 97%. This means that the worst case MMD (Case 1b) is almost as efficient as the MMD under Case 1a. However, examining the bottom-right subplot in each of Figures 4 to 11, shows that the relative efficiencies $RE_{Case\ 1b|Case\ h}$ for $h = 2, 3$ may become rather small when there is a maximum absolute difference between θ_e^U and θ_c^U , either for maximum or minimum values of ϕ^U . For example, a low relative efficiency of about 52% is observed when θ_e^U and θ_c^U are 0.8 and 0.05, respectively (see Figure 5). So we may conclude that using the worst case MMD (Case 1b) instead of MMDs under Case h for $h = 2, 3$ will lead to substantial efficiency loss if θ_e^U and θ_c^U take quite different values.

Figure 6. Efficiency of the MMD under Case 1b compared to the MMD by Case 2 as a function of the parameter ϕ^U for three different sets of cost ratios assuming $\theta = \theta_c^U$ is fixed at 0.20 and θ_e^U varies.

6 Power of the test of no treatment effect

Suppose we want to test whether the intervention is cost-effective or not. To do this we consider the null hypothesis $H_0: \beta_1 = 0$ against the alternative hypothesis $H_A: (b - a)\beta_1 > 0$, that is the intervention is cost-effective. The following test statistic can be constructed²⁷

$$F(1, k - 1; \gamma) = \frac{\hat{\beta}_1^2}{\hat{Var}(\hat{\beta}_1)} \quad (15)$$

Figure 7. Efficiency of the MMD under Case 1b compared to the MMD by Case 2 as a function of the parameter ϕ^U for three different sets of cost ratios assuming $\theta = \theta_c^U$ is fixed at 0.8 and θ_e^U varies.

where $F(1, k-1; \gamma)$ follows the non-central F distribution with the degrees of freedom 1 and $k-1$, respectively, and the non-centrality parameter γ . Note that if H_0 applies, $\gamma = 0$ and if H_A applies, $\gamma > 0$. In the latter case $\gamma = \beta_1^2 \left\{ \text{Var}(\hat{\beta}_1) \right\}^{-1} = \frac{k\beta_1^2}{\sigma_1^2 + \left(\frac{1}{m} + \frac{1}{n}\right)\frac{\sigma^2}{(b-a)^2}} = \frac{kmnES^2}{mnES_V + m + n}$, where $ES = \frac{|b-a|\beta_1}{\sigma}$ is the effect size³⁴ and $ES_V = \frac{(b-a)^2\sigma_1^2}{\sigma^2}$ is the effect size variability,²⁶ which reflect, respectively, the average centre-specific effect size and variability in the centre-specific effect sizes. The power function of the test is $1 - \Pr[F(1, k-1; \gamma) < F_0]$, where F_0 is the $(1 - 2\alpha)$ th percentile of the central $F(1, k-1)$ distribution. For a two-sided test, $1 - 2\alpha$ is replaced by $1 - \alpha$ in calculating F_0 . Note that as m

Figure 8. Efficiency of the *MMD* under *Case 1b* compared to the *MMD* by *Case 3* as a function of the parameter ϕ^U for three different sets of cost ratios and $\rho_{ul}^L = -0.001$, $\rho = -1.0$ assuming $\theta = \theta_e^U$ is fixed at 0.20 and θ_c^U varies.

and n are fixed by the model parameters for optimal and maximin designs, k can be increased by enlarging the budget (see equations (7), (10) to (13)), thus enabling to achieve a certain power level. Upon request, R-codes for calculating optimal and maximin sample sizes that yield sufficient power to test the cost-effectiveness of a treatment can be obtained from the first author.

7 Application in planning a multicentre cost-effectiveness trial: The EVALUATE trial

The optimal design and the *MMD* will be illustrated for a multicentre cost-effectiveness trial known as the EVALUATE trial which compared laparoscopic-assisted hysterectomy (treatment) and

Figure 9. Efficiency of the MMD under Case 1b compared to the MMD by Case 3 as a function of the parameter ϕ^U for three different sets of cost ratios and $\rho_{ul}^L = -0.001$, $\rho = -1.0$ assuming $\theta = \theta_e^U$ is fixed at 0.8 and θ_c^U varies.

standard hysterectomy (control).^{1,2} The data were collected from 25 centres ($k = 25$) with, on the average, $m = 23$ and $n = 11$ patients in each of the treatments per centre. We call this design the actual design ($m_{act} = 23, n_{act} = 11, k_{act} = 25$). The parameter estimates from the EVALUATE trial are $\theta_e = 0.26$, $\theta_c = 0.05$, $\rho_{ul} = 0.66$, $\rho_\varepsilon = -0.083$ and $\phi = 0.60$. Since the parameter σ_c^2 does not play a role in calculating the relative efficiencies in Table 1 and power in Figure 12, the estimate of σ_c^2 from the EVALUATE trial is not reported here. No information on the cost ratios ($c_0/c_1, c_0/c_2$) is available from this study and in order to compute the efficiencies and power of the designs, we considered the cost ratios 2, 5, 10 and 20 used by Raudenbush and Liu,²⁶ assuming equal total budgets for each of the designs.

Figure 10. Efficiency of the MMD under Case 1b compared to the MMD by Case 3 as a function of the parameter ϕ^U for three different sets of cost ratios and $\rho_{u1}^L = -0.001$, $\rho = -1.0$ assuming $\theta = \theta_c^U$ is fixed at 0.20 and θ_e^U varies.

Assuming the parameter estimates from the EVALUATE trial, that is $\Pi_0 = (\theta_e = 0.26, \theta_c = 0.05, \rho_{u1} = 0.66, \rho_\varepsilon = -0.083, \phi = 0.60)$ the efficiency of the actual design (*act*) compared to the optimal design is shown in column 3, Table 1. As can be seen from column 3, Table 1, the optimal design gains efficiency compared to the actual design as the cost ratios in the control arm increase. However, the parameters may not be precisely known at the beginning of the study, so we may not be able to compute the optimal design. One solution could be the use of *MMD* which is based on some boundaries of the plausible ranges for the parameters. The information on the parameters for multicentre cost-effectiveness trials is rather scarce in the literature. Based on the results of the EVALUATE trial, we therefore assume some boundaries for the parameters $\theta = \theta_e^U = \theta_c^U = \max\{\theta_e^U, \theta_c^U\} = 0.40$, $\rho_{u1}^L = \rho_\varepsilon^L = \rho = \min\{\rho_{u1}^L, \rho_\varepsilon^L\} = -0.20$ and $\phi^U = 1$ which

Figure 11. Efficiency of the *MMD* under *Case 1b* compared to the *MMD* by *Case 3* as a function of the parameter ϕ^U for three different sets of cost ratios and $\rho_{ul}^L = -0.001$, $\rho = -1.0$ assuming $\theta = \theta_c^U$ is fixed at 0.8 and θ_e^U varies.

correspond to the *MMD* under *Case 1b* in equation (12) and result in different *MMDs* for different cost ratios. If we assume the parameter estimates from the EVALUATE trial in calculating the optimal design, then the efficiencies of the *MMD* compared to the optimal design lie between 0.73 and 0.83 (see column 4 in Table 1). If the cost ratios increase, the *MMD* loses efficiency compared to the optimal design and in our case, the *MMD* loses up to 27% efficiency (given the assumed values for the cost ratios and the model parameters). Furthermore, if the assumed parameter boundaries used to calculate the *MMD* are the parameter values, that is $\Pi_1 = (\theta_e = 0.40, \theta_c = 0.40, \rho_{ul} = -0.20, \rho_\varepsilon = -0.20, \phi = 1)$, of course the *MMD* becomes optimal compared with the other two designs. The efficiencies of the actual design and the design

Table 1. Relative efficiencies of the actual design of the EVALUATE trial,^{1,2} the maximin design and the optimal design, for the parameters in the EVALUATE trial, (Π_0), and the parameters in Π_1 , specified for different cost ratios.

Cost Ratio (c_0/c_1)	Cost Ratio (c_0/c_2)	$\frac{V(\hat{\beta}_1 \Pi_0,opt)}{V(\hat{\beta}_1 \Pi_0,act)}$	$\frac{V(\hat{\beta}_1 \Pi_0,opt)}{V(\hat{\beta}_1 \Pi_0,MMD)}$	$\frac{V(\hat{\beta}_1 \Pi_1,MMD)}{V(\hat{\beta}_1 \Pi_1,act)}$	$\frac{V(\hat{\beta}_1 \Pi_1,MMD)}{V(\hat{\beta}_1 \Pi_1,opt)}$
2	2	0.7878	0.8289	0.3451	0.7238
2	5	0.6859	0.8085	0.3300	0.7148
2	10	0.6090	0.7964	0.3107	0.7113
2	20	0.5453	0.7870	0.2921	0.7095
5	2	0.9084	0.8085	0.4369	0.7148
5	5	0.8569	0.7831	0.4685	0.7091
5	10	0.7812	0.7681	0.4651	0.7089
5	20	0.7018	0.7566	0.4491	0.7106
10	2	0.9412	0.7964	0.4802	0.7113
10	5	0.9516	0.7681	0.5666	0.7089
10	10	0.9001	0.7517	0.5957	0.7120
10	20	0.8230	0.7393	0.5965	0.7173
20	2	0.9302	0.7870	0.4982	0.7095
20	5	0.9921	0.7566	0.6349	0.7102
20	10	0.9749	0.7393	0.7067	0.7173
20	20	0.9125	0.7269	0.7384	0.7267

Note. Maximin design (*MMD*) is always based on the parameter vector Π_1 , *act* stands for the actual design of the EVALUATE trial and the optimal design (*opt*) is based on the parameter estimates as found in the EVALUATE trial, Π_0 .

which is optimal for Π_0 compared to the efficiency of the *MMD* are shown in columns 5 and 6 of Table 1, respectively.

Next, we compared power among the optimal design, the *MMD* and the actual design of the EVALUATE trial when varying the effect size (*ES*) and its variability (ES_V). Similar to Raudenbush and Liu²⁶ we used the values 0.05, 0.10 and 0.15 for ES_V , which are considered to reflect small, medium and large effect size variability, and additionally, we considered three more values for ES_V , namely 0.025, 0.20, 0.25. Figure 12 displays power for the optimal, maximin and actual design of the EVALUATE trial indicated by different line types for a given set of cost ratios ($c_0/c_1 = 2$, $c_0/c_2 = 2$) where the effect size (*ES*) and the effect size variability (ES_V) vary within and between plots, respectively. As can be seen in Figure 12, for all designs, an increasing effect size (*ES*) leads to larger power and an increasing effect size variability (ES_V) leads to smaller power. It can be shown that the optimal design loses more power than the *MMD* as the effect size variability (ES_V) increases. As a result, the optimal design may have less power than the *MMD* when the effect size variability (ES_V) is high (as illustrated by the bottom-right subplot of Figure 12). This means that if we have large effect size variability (ES_V), the *MMD* may become better than the optimal design. Further, for other sets of cost ratios the patterns emerging from the plots were similar and for that reason these results were omitted.

8 Discussion and conclusions

This paper provides optimal and maximin sample size formulas for multicentre cost-effectiveness trials that evaluate the cost-effectiveness of interventions on a continuous scale. In such trials the cost-effectiveness of interventions may vary markedly between centres and we addressed this between-centre variability of outcomes in deriving sample size formulas. The optimality criterion

Figure 12. Statistical power for the optimal design, the *MMD* and the actual design of the EVALUATE trial for a given set of cost ratios $c_0/c_1 = 2$, $c_0/c_2 = 2$.

which minimizes the variance of the treatment effect estimator for a given constraint regarding the cost for selecting samples was used for the derivation of sample size formulas. The optimal sample size formulas were obtained for a general treatment coding scheme. The optimal sample size formulas in equation (7) depend on six model parameters. The researcher should have prior knowledge on these parameters to obtain the optimal sample sizes but the values of these parameters are often unknown in the design stage of a trial. In response to this, we proposed the maximin sample sizes for multicentre cost-effectiveness trials to accommodate the uncertainty regarding the model parameters. These maximin sample sizes optimize the worst-case scenario in terms of model parameters and ensure a desired level of power for prespecified ranges of parameter values. Four different maximin sample size formulas in equations (10) to (13) were derived based on the signs of the lower bounds for the correlation between the random slope for costs and effects, ρ_{u1} , and the correlation between random costs and effects at the individual level, ρ_ε . The *MMD* in

equation (11) is the worst case compared to the *MMDs* in equations (10), (12) and (13). In case of insufficient knowledge on the signs of the lower bounds for the correlations ρ_{u1} and ρ_{ε} , we may start with worst case *MMD* in equation (11), in line with the maximin principle. We studied numerically the efficiency of the worst case *MMD* instead of using the *MMDs* in equations (10), (12) and (13) and the numerical evaluation shows that the worst case *MMD* and the *MMD* in equation (10) are almost equally efficient. On the other hand, the worst case *MMD* may lose much efficiency compared to using the *MMDs* in equations (12) and (13). The efficiency loss is larger for larger differences between the upper bounds of θ_e and θ_c . In our numerical evaluation, the maximum efficiency loss was found to be 48%, when $\theta_e^U = 0.8$ and $\theta_c^U = 0.05$ (see Figure 5). Furthermore, we derived an expression how the optimal or maximin sample sizes for multicentre cost-effectiveness trials are related to the power of the test on cost-effectiveness, effect size and effect size variability. This expression can be used to obtain the required budget in calculating optimal or maximin sample sizes for a given desired level of power for the test of cost-effectiveness of a treatment.

This study has some limitations. First, the sample size formulas are true for equal numbers of individuals across centres in each of the treatment conditions but in practice the number of individuals per treatment arm may vary from one centre to the other. In this case one can use correction methods as suggested by Van Breukelen et al.³⁵ and Candel et al.³⁶ to adjust the sample sizes for variation in centre sizes. Second, the random-effects model in equation (1) assumes a random interaction effect between treatment and centre. However, the cost-effectiveness of interventions may differ from one clinical trial to another. This means that there may also be a random treatment-by-trial interaction.⁴ In this sense the current random-effects model is limited in terms of its ability to generalize. However, we consider the model more adequate than a fixed-effects approach, where conclusions are strictly limited to the specific centres examined, even if they are randomly selected. Third, the *NMB* in equation (2) follows a normal distribution because the costs and effects in equation (1) are assumed to follow a normal distribution. Cost data are usually positively skewed. Some reviews have suggested that the analysis of cost-effectiveness data is robust against violation of the normality assumption.^{37,38} Further work however is needed to test the generalizability of such robustness in different contexts especially for small samples and highly positively skewed outcome measures.

Funding

This research was supported by the Netherlands Organisation for Health Research and Development (ZonMw) [Grant Number 80-82310-97-12201].

Conflict of interest

None declared.

References

1. Sculpher M, Manca A, Abbott J, et al. The cost-effectiveness of laparoscopic-assisted hysterectomy in comparison with standard hysterectomy: results from a randomised trial. *Br Med J* 2004; **328**: 134–139.
2. Manca A, Rice N, Sculpher MJ, et al. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Econ* 2005; **14**: 471–485.
3. Fedorov V and Jones B. The design of multicentre trials. *Stat Methods Med Res* 2005; **14**: 205–248.
4. Senn SJ. The many modes of meta. *Drug Inform J* 2000; **34**: 535–549.
5. Senn SJ. Some controversies in planning and analyzing multi-centre trials. *Stat Med* 1998; **17**: 1753–1765.
6. Dragalin V, Fedorov V, Jones B, et al. Estimation of the combined response to treatment in multicenter trials. *J Biopharm Stat* 2001; **11**: 275–295.

7. Kahan BC and Morris TP. Analysis of multicentre trials with continuous outcomes: when and how should we account for centre effects? *Stat Med* 2013; **32**: 1136–1149.
8. Gallo PP. Center-weighting issues in multicenter clinical trials. *J Biopharm Stat* 2000; **10**: 145–163.
9. Senn SJ. A note regarding 'random effects'. *Stat Med* 2014; **33**: 2876–2877.
10. Grieve R, Nixon R, Thompson SG, et al. Multilevel models for estimating incremental net benefits in multinational studies. *Health Econ* 2007; **16**: 815–826.
11. Carey K. A multi-level modelling approach to analysis of patient costs under managed care. *Health Econ* 2000; **9**: 435–446.
12. Grieve R, Nixon R, Thompson SG, et al. Using multilevel models for assessing the variability of multinational resource use and cost data. *Health Econ* 2005; **14**: 185–196.
13. Nixon RM and Thompson SG. Incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. *Health Econ* 2005; **14**: 1217–1229.
14. Rice N and Jones A. Multilevel models and health economics. *Health Econ* 1997; **6**: 561–575.
15. Thompson SG, Nixon RM and Grieve R. Addressing the issues that arise in analyzing multicentre cost data, with application to a multinational study. *Health Econ* 2006; **25**: 1015–1028.
16. Willan AR, Pinto EM, O'Brien BJ, et al. Country specific cost comparisons from multinational clinical trials using empirical Bayesian shrinkage estimation: the Canadian ASSENT-3 economic analysis. *Health Econ* 2005; **14**: 327–338.
17. Willan AR, Pinto EM and O'Brien BJ. Cost-effectiveness analysis for multinational clinical trials. *Stat Med* 2005; **24**: 1965–1982.
18. Verbeke G and Molenberghs G. *Linear mixed models for longitudinal data*. New York: Springer, 2000.
19. Al MJ, Van Hout BA, Michel BC, et al. Sample size calculation in economic evaluations. *Health Econ* 1998; **7**: 327–335.
20. Briggs A and Gray AM. Power and sample size calculations for stochastic cost-effectiveness analysis. *Med Decis Making* 1998; **18**: 81–92.
21. Laska EM, Meisner M and Siegel C. Power and sample size in cost-effectiveness analysis. *Med Decis Making* 1999; **19**: 339–343.
22. Zethraeus N, Johannesson M, Jönsson B, et al. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. *Pharmacoeconomics* 2003; **21**: 39–48.
23. Willan AR. Sample size determination for cost-effectiveness trials. *Pharmacoeconomics* 2011; **29**: 933–949.
24. Moerbeek M, Van Breukelen GJP and Berger MPF. Design issues for experiments in multilevel population. *J Educ Behav Stat* 2000; **25**: 271–284.
25. Moerbeek M, Van Breukelen GJP and Berger MPF. A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *J Clin Epidemiol* 2003; **56**: 341–350.
26. Raudenbush ST and Liu X. Statistical power and optimal design for multisite randomized trials. *Psychol Methods* 2000; **5**: 199–213.
27. Liu X. Statistical power and optimum sample allocation ratio for treatment and control having unequal costs per unit of randomization. *J Educ Behav Stat* 2003; **28**: 231–248.
28. Berger MPF and Wong WK. *An introduction to optimal designs for social and bio-medical research*. Chichester: Wiley, 2009.
29. Atkinson AC, Donev AN and Tobias RD. *Optimum experimental designs, with SAS*. Oxford: Oxford university press, 2007.
30. Wu CFJ. Efficient sequential designs with binary data. *J Am Stat Assoc* 1985; **80**: 974–984.
31. Spiegelhalter DJ, Abrams KR and Myles JP. *Bayesian approaches to clinical trials and health-care evaluation*. Chichester: Wiley, 2007.
32. Stinnett A and Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making* 1998; **18**: S68–S80.
33. Moerbeek M. Randomization of clusters versus randomization of persons within clusters: which is preferable? *Am Stat* 2005; **59**: 173–179.
34. Cohen J. A power primer. *Psychol Bull* 1992; **112**: 155–159.
35. Van Breukelen GJP, Candel MJJM and Berger MPF. Relative efficiency of unequal versus equal cluster sizes in cluster randomized and multicentre trials. *Stat Med* 2007; **26**: 2589–2603.
36. Candel MJJM, Van Breukelen GJP, Kotova L, et al. Optimality of unequal cluster sizes in multilevel studies with small sample sizes. *Commun Stat Simulat Comput* 2008; **37**: 222–239.
37. Willan AR, Briggs AH and Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004; **13**: 461–475.
38. Thompson SG and Barber JA. How should cost data in pragmatic randomised trials be analysed? *Br Med J* 2000; **320**: 1197–1200.

Appendix I: Derivation of the maximin sample sizes and its MMV

The proof involves the following steps: First, show for which boundary values of the parameters, the efficiency (Eff) in equation (9) of the main text becomes minimum. Second, show which values of m , n and k optimize the minimum Eff .

Step 1: The efficiency of the treatment effect estimator can be written as

$$Eff = \left\{ Var(\hat{\beta}_1) \right\}^{-1} = k \left\{ \sigma_1^2 + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^2}{(b-a)^2} \right\}^{-1} \quad (16)$$

where $\sigma_1^2 = (\phi\theta_e + \theta_c - 2\rho_{u1}\sqrt{\theta_e\theta_c\phi})\sigma_c^2$ and $\sigma^2 = (b-a)^2 \left\{ \phi(1-\theta_e) + (1-\theta_c) - 2\rho_\varepsilon\sqrt{(1-\theta_e)(1-\theta_c)} \right\} \sqrt{\phi}\sigma_c^2$. As is evident from equation (16), to find the minimum Eff the upper bound of σ_c^2 , that is σ_c^{2U} , should be chosen (U stands for upper boundary). The partial derivatives of equation (16) with respect to ρ_{u1} and ρ_ε then become $\frac{\partial Eff}{\partial \rho_{u1}} = \frac{2Eff^2\sigma_c^{2U}\sqrt{\theta_e\theta_c\phi}}{k} \geq 0$ and $\frac{\partial Eff}{\partial \rho_\varepsilon} = \frac{2Eff^2\sigma_c^{2U}}{k} \left(\frac{1}{m} + \frac{1}{n} \right) \sqrt{(1-\theta_c)(1-\theta_e)\phi} \geq 0$, respectively. These partial derivatives show that the efficiency always goes up as ρ_{u1} and ρ_ε increase. So the minimum efficiency occurs at the lower boundaries of ρ_{u1} and ρ_ε . Let the lower bounds of ρ_{u1} and ρ_ε be ρ_{u1}^L ($-1 \leq \rho_{u1}^L \leq 1$) and ρ_ε^L ($-1 \leq \rho_\varepsilon^L \leq 1$) (L stands for lower boundary). So the minimum efficiency with respect to ρ_{u1} and ρ_ε becomes

$$Eff^* = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (17)$$

where $\sigma_1^{*2} = (\phi\theta_e + \theta_c - 2\rho_{u1}^L\sqrt{\theta_e\theta_c\phi})\sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2 \left\{ \phi(1-\theta_e) + (1-\theta_c) - 2\rho_\varepsilon^L\sqrt{(1-\theta_e)(1-\theta_c)} \right\} \sqrt{\phi}\sigma_c^{2U}$. Since ρ_{u1}^L and ρ_ε^L take values between -1 and $+1$, the following cases are possible:

Case 1a ($-1 \leq \rho_{u1}^L < \rho_\varepsilon^L < 0$): Taking the partial derivative of (17) with respect to ϕ gives $\frac{\partial Eff^*}{\partial \phi} = -\frac{Eff^{*2}\sigma_c^{2U}}{k} \left[\left(\theta_e - \rho_{u1}^L\sqrt{\frac{\theta_e\theta_c}{\phi}} \right) + \left(\frac{1}{m} + \frac{1}{n} \right) \left\{ (1-\theta_e) - \rho_\varepsilon^L\sqrt{\frac{(1-\theta_e)(1-\theta_c)}{\phi}} \right\} \right] \leq 0$ implying that Eff^* decreases with increasing ϕ . So Eff^* becomes minimum at the upper bound of ϕ which is, say ϕ^U ($0 \leq \phi^U < \infty$). Now we partially differentiate Eff^* with respect to θ_e : $\frac{\partial Eff^*}{\partial \theta_e} = \frac{-Eff^{*2}\sigma_c^{2U}}{k} \left[\phi \left\{ 1 - \left(\frac{1}{m} + \frac{1}{n} \right) \right\} - \rho_{u1}^L\sqrt{\frac{\theta_c\phi}{\theta_e}} + \rho_\varepsilon^L \left(\frac{1}{m} + \frac{1}{n} \right) \sqrt{\frac{(1-\theta_c)\phi}{(1-\theta_e)}} \right]$. Since $\left(\frac{1}{m} + \frac{1}{n} \right) \leq 1$ for all $m \geq 2$ and $n \geq 2$, the derivative $\frac{\partial Eff^*}{\partial \theta_e} \leq 0$ if $\left(\sqrt{\frac{\theta_c}{\theta_e}} - \sqrt{\frac{1-\theta_c}{1-\theta_e}} \right) \geq 0$, which is equivalent to $\theta_c \geq \theta_e$. Similarly, $\frac{\partial Eff^*}{\partial \theta_c} \leq 0$, if $\theta_e \geq \theta_c$. So the efficiency will be minimum for both θ_e and θ_c if $\theta_e = \theta_c = \max\{\theta_e^U, \theta_c^U\} = \theta$, where $\theta_e \in [\theta_e^L, \theta_e^U]$ with $0 < \theta_e^L \leq \theta_e^U < 1$ and $\theta_c \in [\theta_c^L, \theta_c^U]$ with $0 < \theta_c^L \leq \theta_c^U < 1$ are the plausible ranges for the parameters θ_e and θ_c respectively. Note that 0 and 1 were excluded from the boundaries of θ_e and θ_c , because we could not get any solution for the maximin sample sizes in these cases. So the minimum efficiency with respect to the model parameters is given by

$$Eff^{\min} = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (18)$$

where $\sigma_1^{*2} = \theta(\phi^U + 1 - 2\rho_{u1}^L\sqrt{\phi^U})\sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2(1-\theta)(\phi^U + 1 - 2\rho_\varepsilon^L\sqrt{\phi^U})\sigma_c^{2U}$.

Case 1b ($-1 \leq \rho_\varepsilon^L \leq \rho_{u1}^L < 0$): Then the derivative $\frac{\partial Eff^*}{\partial \phi} \leq 0$. On the other hand, the minimum efficiency with respect to θ_e and θ_c also depends on the value of ρ_{u1}^L and ρ_ε^L in this case. In order to find the values for θ_e and θ_c , which minimize the efficiency in equation (17), it is necessary to assume that $\rho_{u1}^L = \rho_\varepsilon^L = \min\{\rho_{u1}^L, \rho_\varepsilon^L\} = \rho$. However, this choice minimizes the efficiency more than needed when $\rho_\varepsilon^L < \rho_{u1}^L$ but ensures a certain level of power also in this case. Now the derivatives $\frac{\partial Eff^*}{\partial \theta_e} \leq 0$ and $\frac{\partial Eff^*}{\partial \theta_c} \leq 0$ if $\theta_e = \theta_c = \max\{\theta_e^U, \theta_c^U\} = \theta$. So the minimum efficiency becomes

$$Eff^{\min} = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (19)$$

where $\sigma_1^{*2} = \theta(\phi^U + 1 - 2\rho\sqrt{\phi^U})\sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2(1-\theta)(\phi^U + 1 - 2\rho\sqrt{\phi^U})\sigma_c^{2U}$.

Case 2 ($\rho_{u1}^L \geq 0$ and $\rho_\varepsilon^L \geq 0$): In this case, the lower bound for the efficiency with respect to ρ_{u1} and ρ_ε is obtained by setting $\rho_{u1} = \rho_\varepsilon = 0$ in equation (17) which is the following

$$Eff^* = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (20)$$

where $\sigma_1^{*2} = (\phi\theta_e + \theta_c)\sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2\{\phi(1-\theta_e) + (1-\theta_c)\}\sigma_c^{2U}$. Then the derivatives are $\frac{\partial Eff^*}{\partial \phi} = -\frac{Eff^{*2}\sigma_c^{2U}}{k}\left\{\theta_e + \left(\frac{1}{m} + \frac{1}{n}\right)(1-\theta_e)\right\}$, $\frac{\partial Eff^*}{\partial \theta_e} = -\frac{Eff^{*2}\sigma_c^{2U}\phi}{k}\left\{1 - \left(\frac{1}{m} + \frac{1}{n}\right)\right\}$ and $\frac{\partial Eff^*}{\partial \theta_c} = -\frac{Eff^{*2}\sigma_c^{2U}}{k}\left\{1 - \left(\frac{1}{m} + \frac{1}{n}\right)\right\}$, which also satisfy $\frac{\partial Eff^*}{\partial \phi} \leq 0$, $\frac{\partial Eff^*}{\partial \theta_e} \leq 0$ and $\frac{\partial Eff^*}{\partial \theta_c} \leq 0$. This shows equation (20) reaches its minimum at $\phi = \phi^U$, $\theta_e = \theta_e^U$ and $\theta_c = \theta_c^U$ where $0 \leq \phi^U < \infty$, $0 \leq \theta_e^U \leq 1$ and $0 \leq \theta_c^U \leq 1$, which is given by

$$Eff^{\min} = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (21)$$

where $\sigma_1^{*2} = (\phi^U\theta_e^U + \theta_c^U)\sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2\{\phi^U(1-\theta_e^U) + (1-\theta_c^U)\}\sigma_c^{2U}$.

Case 3 ($\rho_{u1}^L < 0$ and $\rho_\varepsilon^L \geq 0$): We get the minimum efficiency with respect to ρ_{u1} and ρ_ε from equation (17), which is

$$Eff^* = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (22)$$

where $\sigma_1^{*2} = (\phi\theta_e + \theta_c - 2\rho_{u1}^L\sqrt{\theta_e\theta_c\phi})\sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2\{\phi(1-\theta_e) + (1-\theta_c)\}\sigma_c^{2U}$. Now the derivatives of equation (22) with respect to ϕ , θ_e and θ_c are

$$\begin{aligned} \frac{\partial Eff^*}{\partial \phi} &= -\frac{Eff^{*2}\sigma_c^{2U}}{k} \left\{ \theta_e - \rho_{u1}^L \sqrt{\frac{\theta_e\theta_c}{\phi}} + \left(\frac{1}{m} + \frac{1}{n} \right) (1-\theta_e) \right\}, \\ \frac{\partial Eff^*}{\partial \theta_e} &= \frac{-Eff^{*2}\sigma_c^{2U}\phi}{k} \left\{ 1 - \left(\frac{1}{m} + \frac{1}{n} \right) - \rho_{u1}^L \sqrt{\frac{\theta_c}{\phi\theta_e}} \right\} \text{ and} \end{aligned}$$

$\frac{\partial Eff^*}{\partial \theta_c} = \frac{-Eff^{*2}\sigma_c^{2U}\phi}{k} \left\{ 1 - \left(\frac{1}{m} + \frac{1}{n} \right) - \rho_{u1}^L \sqrt{\frac{\theta_e}{\phi\theta_c}} \right\}$ and these derivatives satisfy $\frac{\partial Eff^*}{\partial \phi} \leq 0$, $\frac{\partial Eff^*}{\partial \theta_e} \leq 0$ and $\frac{\partial Eff^*}{\partial \theta_c} \leq 0$.

The minimum efficiency occurs at $\phi = \phi^U$, $\theta_e = \theta_e^U$ and $\theta_c = \theta_c^U$ where $0 \leq \phi^U < \infty$, $0 \leq \theta_e^U \leq 1$ and $0 \leq \theta_c^U \leq 1$ and is equal to

$$Eff^{\min} = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (23)$$

where $\sigma_1^{*2} = (\phi^U \theta_e^U + \theta_c^U - 2\rho_{u1}^L \sqrt{\theta_e^U \theta_c^U \phi^U}) \sigma_c^{2U}$ and $\sigma^{*2} = (b-a)^2 \{ \phi^U (1 - \theta_e^U) + (1 - \theta_c^U) \} \sigma_c^{2U}$.

Case 4 ($\rho_{u1}^L \geq 0$ and $\rho_\varepsilon^L < 0$): The minimum efficiency in this case with respect to ρ_{u1} and ρ_ε is

$$Eff^* = k \left\{ \sigma_1^{*2} + \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\sigma^{*2}}{(b-a)^2} \right\}^{-1} \quad (24)$$

where $\sigma_1^{*2} = (\phi \theta_e + \theta_c) \sigma_c^{2U}$, $\sigma^{*2} = (b-a)^2 \{ \phi(1 - \theta_e) + (1 - \theta_c) - 2\rho_\varepsilon^L \sqrt{(1 - \theta_e)(1 - \theta_c)\phi} \} \sigma_c^{2U}$. Now $\frac{\partial Eff^*}{\partial \phi} \leq 0$ and Eff^* becomes minimum at $\phi = \phi^U$. The partial derivative of equation (24) with respect to θ_e is $\frac{\partial Eff^*}{\partial \theta_e} = \frac{-Eff^{*2} \sigma_c^{2U}}{k} \left[\phi \left\{ 1 - \left(\frac{1}{m} + \frac{1}{n} \right) \right\} + \rho_\varepsilon^L \left(\frac{1}{m} + \frac{1}{n} \right) \sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \right]$ and $\frac{\partial Eff^*}{\partial \theta_e} \leq 0$ if $\left[\phi \left\{ 1 - \left(\frac{1}{m} + \frac{1}{n} \right) \right\} + \rho_\varepsilon^L \left(\frac{1}{m} + \frac{1}{n} \right) \sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \right] \geq 0$, or $\sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \geq \rho^* A$ where $A = \frac{m+n}{mn-m-n}$ and $\rho_\varepsilon^L = -\rho^* (0 < \rho^* \leq 1)$. However, $\frac{\partial Eff^*}{\partial \theta_c} \leq 0$ if $\sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \leq \frac{1}{\rho^* A}$. This means that the Eff^* will decrease with respect to θ_e and θ_c if $\sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \geq \rho^* A$ and $\sqrt{\frac{(1 - \theta_e)\phi}{(1 - \theta_c)}} \leq \frac{1}{\rho^* A}$, respectively. However, these two conditions in general are not satisfied at the same time. Hence, this derivation does not lead to a simple solution for the maximin sample sizes in Step 2.

Step 2: The maximin sample sizes, and the optimal variance (MMV) of these sample sizes, in each of the above cases (except *Case 4*) are obtained by maximizing the $Eff^{\min}$ in each of the corresponding cases for the budget constraint in equation (6) of the main text. The resulting maximin sample sizes and MMV s are shown in equations (10) to (13) of the main text.