Revisiting the Estimation of Multilevel Modeling and Power Analysis for Multisite Randomized Trials

Zhenqiu Lu0000-0001-9482-1368 \*†   
† Quantitative Methodologies, University of Georgia, Athens, 30602, GA, USA  
\* Corresponding author. Email: zlu@uga.edu

|  |
| --- |
| **Abstract** Challenges arise in multisite randomized trials in power analysis and optimal design due to the complexities of random coefficients and hierarchical structures. To address these challenges, this article revisits the estimation of multilevel models (MLM). It presents various approaches to deriving parameter estimates, incorporating different algebraic derivations—basic or matrix algebra, along with different designs—balanced and unbalanced designs, different coding schemes for treatment and control groups, and some specific special cases. By showing the similarities, differences, and relationships among these estimates, the study provides a clearer framework, enhances transparency, and improves interpretability. Furthermore, by comparing the derived results with existing literature, it establishes connections that deepen researchers' understanding of MLM estimation in study design and power analysis. |

**Keywords:**  multilevel model estimation, multisite randomized trials, balanced and unbalanced design, power analysis

**1. Introduction**

Power analysis (e.g., Cohen, 1988) is essential in randomized trials studies, particularly in determining the sample size to detect treatment effects. An optimal sample allocation (Raudenbush & Liu, 2000) is to maximize power while minimizing sampling costs. In power analysis and optimal design, the test statistic for the main effect of treatment follows a non-central t-distribution under the alternative hypothesis. Since power is positively related to the non-centrality parameter, maximizing power requires maximizing this parameter, which means minimizing the variance of the treatment effect estimate.

Despite its importance, several challenges exist. First, in multisite randomized trials, the incorporated random coefficients add complexity to estimation. Hierarchical linear models (HLM; e.g., Raudenbush & Bryk, 2002), also known as multilevel models (MLM; e.g., Snijders, & Bosker, 2012) or mixed-effect models (e.g., Fitzmaurice, Laird & Ware, 2012), have various estimation methods, such as full maximum likelihood estimation (FML; Goldstein, 1986; Longford, 1987), restricted/residual maximum likelihood estimation (RML; Mason et al., 1983; Raudenbush & Bryk, 1986), least square estimation (LS), exploratory estimation methods and others. Additionally, estimations can be derived using either basic algebra or matrix algebra, leading to discrepancies in formats. Second, most existing derivation in multisite randomized trials assume balanced designs (e.g., Raudenbush & Bryk, 2002; Dong, Kelcey & Spybrook, 2021), where control and experimental groups have equal sample sizes. However, real-world studies often involve unbalanced designs, where sample sizes differ in control and experimental groups. Assuming balanced designs limits their applicability to practical scenarios. Third, different coding schemes for treatment and control groups create inconsistencies in estimation results. Some assign 0 to the control group and 1 to the experimental group, while others use -0.5 and +0.5, respectively. Many existing power analysis and optimal design results (e.g., Raudenbush & Liu, 2000; Dong, Kelcey & Spybrook, 2021) are derived under the latter coding scheme in balanced designs, limiting their generalizability to other coding schemes and unbalanced designs.

To address these challenges, this article aims to (1) provide a clear and systematic approach to deriving MLM parameter estimates and related results for mixed-effect models in multisite randomized trials to enhance clarity and interpretability, (2) extend estimation results to more general cases, particularly unbalanced designs, making it more applicable to real-world research, while also deriving balanced design results as a special case, (3) provide estimation approaches for different coding schemes to ensure transparency in their differences and implications, (4) compare the derived results with existing literature and establish connections among them, (5) help researchers to improve the understanding of theoretical foundation of MLM estimation in multisite randomized trials.

This paper is organized as follows. Section 2 reviews estimation in the general multilevel model using matrix algebra. Section 3 focuses on estimation in multisite randomized trials, using individual-level formulations and basic algebra. It also compares estimates derived from different coding schemes, study designs, and special cases. Section 4 applies the results to power analysis, and Section 5 concludes the paper.

**2. The general MLM model**

Suppose there are *J* sites and is the sample size at the site. The total sample size N is . Let be the observed data point for object at site, and be an vector of variables . In a two-level hierarchical analysis, the general MLM has two levels (Raudenbush & Bryk, 2002). At Level 1, the model is

with (1)

where is an design matrix of independent variables at level 1, is the number of independent variables, is a *(* vector of level 1 random parameters, is an identity matrix, and is an vector of normally distributed random errors with mean **0** and covariance matrix , which is usually assumed for simplicity. At Level 2, the model is

with (2)

where is a matrix of level 2 predictors, is an vector of fixed effects, is a vector of level-2 random residuals assumed to follow a multivariate normal distribution with mean 0 and covariance matrix of dimension . By combining (1) and (2), the composite form is

***2. 1 The estimate of the random coefficient at level 1***

To estimate , we apply the ordinary least-squares (OLS) on regression and obtain:

(3)

*2.1.1 Conditional expectation and conditional variance*

The conditional expectation of is

(4)

Assuming a known variance , the conditional variance of is

(5)

For simplicity, we use to denote .

*2.1.2 Unconditional expectation and unconditional variance*

From equations (1), (2) and (3), we obtain

where with . Thus, the unconditional expectation of is

(6)

and the unconditional variance of , assuming known variance and , is given by

(7)

***2. 2 The estimate of the fixed effect at level 2***

Since where and , we apply the generalized least-squares (GLS) to estimate and obtain:

(8)

The expectation of is

and the variance of is

(9)

***2. 3 The estimate of random components at both levels***

When the residual variance at level 1 and the residual covariance matrix at level 2 are unknown, there are no closed-form mathematical formulas available to estimate them. Instead, iterative numerical procedures are used to obtain estimates. Several iterative estimation methods exist, including iterative GLS and iterative MLE.

For the iterative GLS, we first apply the OLS method to the composite model to obtain initial parameter estimates and an estimate of the residual covariance matrix. Then re-fit the composite model using GLS, treating the estimated residual covariance matrix as if it were the true residual covariance matrix. This yields updated estimates of the fixed effects parameters and residual covariance matrix. We repeat the previous step iteratively until the results converges.

For the iterative MLE, we first assign initial values to the unknown parameters. Then compute the likelihood function based on these parameter values, and update the parameter estimates based on the current likelihood function. We repeat the previous step multiple times until the difference between successive maximum likelihood values is less than a given criterion (e.g., 0.000001), indicating convergence. The final parameter estimates corresponding to the maximum likelihood function are the iterative ML estimates. There are two types of MLEs: FML and RML. While FML provides biased estimates of variance components, RML produces unbiased variance component estimates.

**3. The mixed-effect model in multisite randomized trials**

In multisite randomized trials, it is typically assumed that each site includes both control and experimental groups. Considering the simplest case with only a grouping variable X and no additional covariates, the Level 1 model is given by

*with*  (10)

assuming that both the control and experimental groups share the same residual variance . If for the control group (C) and for the experimental/treatment group (E), then represents the average outcome of the control group at site, while represents the average difference between the treatment and control groups at site . The Level 2 model is specified as:

and with (11)

where is the mean of the control group across all *J* sites, and is the mean difference between the treatment and control group across *J* sites, is the variance of control group means, is the variance of mean difference between the treatment and control groups, and their covariance . By combining (10) and (11), we obtain the composite model:

By comparing this model and the general MLM model, (1) - (2) reduce to (10) - (11), respectively, by setting, , and .

Note that in multisite randomized trials, study designs can be balanced or unbalanced, depending on sample size allocation. Let be the sample size at site , with and the sample sizes of the control and experimental groups, respectively, such that . In a balanced design, both groups have equal sample sizes, i.e., . In contrast, an unbalanced design allows the control and experiment groups to have different sample sizes. If further assume as the proportion of participants assigned to the experimental group, then and .

***3. 1 The estimate of the random coefficient and at level 1***

If we use dummy coding for the grouping variable *X* by setting for the control group and for the experimental group, then and can be estimated as

where and are the sample means of the control and experimental groups at site , respectively. Since and are random-effect estimates, they have both conditional and unconditional expectations and variances. By comparing these estimates with (3), the coding of X matters. Different values of X lead to variations in the form of , especially for , resulting in different interpretations.

*3.1.1 Conditional expectation and conditional variance of the estimates*

The conditional expectations of and are

By applying the central limit theorem (CLT), the conditional variances of and are

(12)

(13)

*3.1.2 Unconditional expectation and unconditional variance of the estimates*

By the law of total expectation , the unconditional expectations of and are

By the law of total variance , the unconditional variances of and are

(14)

(15)

In the special case where for all sites, the total sample size is , with and . Under this condition, the variances in (14) and (15) simplify to

*3.1.3 Different coding schemes of X in a balanced design*

Different codings of the group variable X lead to variations in the form of in equation (3), particularly , resulting in different interpretations. In addition to dummy coding, another commonly used alternative is effect coding, where for the control group and for the experimental group (sometimes rescaled to -1/+1). This coding scheme is most appropriate in balanced designs, where the sample sizes in the control and experimental groups are equal (i.e., ). In this case, the intercept represents the grand mean of *Y* at site , and continues to represent the treatment effect at that site. When the design is unbalanced, the intercept becomes a weighted grand mean, making the interpretation of coefficients less intuitive. Therefore, under the coding in a balanced design, the coefficients can be estimated as:

Under this condition, the variances in (12) - (15) become:

(16)

(17)

In the special case where for all sites, these variances simplify to

(18)

(19)

(20)

(21)

The results in (18) - (19) are identical to those in equation (6) on page 201 of Raudenbush & Liu (2000), and (20) - (21) are identical to those in equation (7) on the same page. It is important to note that these results apply specifically to the coding scheme of  */* in a balanced design. When using coding, whether in a balanced or unbalanced design, the variances are given by (12)-(15).

***3. 2 The estimate of the fixed-effect and*** ***at level 2***

If for the control group and for the experimental group, then and can be estimated as

(22)

(23)

The expectation of the and are

By applying (14) and (15), the variances of and are

(24)

(25)

In the special case where for all sites with and , equations (22) - (25) simplify to

In a balanced design and for the control group and for the experimental group, is the expected value across all groups, while is the expected mean difference between the experimental and control groups. In this case, and can be estimated as follows:

By applying (16) and (17), the variances of the estimates and are

(26)

(27)

In the special case where for all sites, the estimates and their variances in equations (26)-(27) simplify to

(28)

(29)

The results (28) - (29) match those in equation (13) on page 202 of Raudenbush & Liu (2000). Again, these results specifically apply to the / coding in a balanced design. For the coding, whether in a balanced or unbalanced design, the variances are given by (24)-(25).

***3. 3 The estimate of random components at both levels***

When the level 1 random component and the level 2 random component are unknown, iterative algorithms are used for estimation. Suppose the Expectation-Maximization (EM) algorithm is applied. The E-step and M-step are detailed as follows.

E-step: The composite model can be expressed in matrix form as

where is an vector of , is an design matrix, is a vector of fixed-effect parameters,is a vector of level 2 randomness, and is an vector of residues. Based on the assumptions in (1) and (2), the joint distribution of and is given by

Therefore, the distribution of conditional on is

M-step: Once the random are generated, the parameters **,**  and can be estimated by maximizing the likelihood or using least-square estimation methods.

Specifically, given the random coefficient estimates and , the estimates and are

(30)

(31)

where

The results (30)-(31) are derived from (14)-(15) and

When the design is balanced, if we assume and the coding is /, then (30) - (31) simply to

(32)

(33)

The results (32) - (33) align with Equation (12) on page 202 of Raudenbush & Liu (2000).

**4. Power analysis**

With the derived parameter estimates, now we can compute the statistical power for testing the average treatment effect. In power analysis, the t statistic for the main effect of treatment with degrees of freedom *J-1* is calculated as follows:

For a non-zero treatment effect under the alternative hypothesis, the *t* statistic follows a non-central *t* distribution with the degrees of freedom *J-1* and a non-centrality parameter .

In an unbalanced design,

If the type-1 error is , then the critical value for a two-sided *t* test is , and the , and in a one-sided t test, , where is the cumulative distribution function (CDF) of a non-central *t* distribution.

In a balanced design, the simplify to

In a two-sided *t* test, , and in a one-sided *t* test, . In the special case where for all sites, the non-centrality parameter becomes

This corresponds to the square root of the in Equation (15) on page 202 of Raudenbush & Liu (2000), where they used a non-central *F* distribution instead of a non-central *t* distribution. Their λ parameter was derived under the assumption of in a balanced design, which is a special case of the more general scenarios developed in this article.

**5. Discussion**

This article revisits the estimation of multilevel models in power analysis and optimal design, with a focus on mixed-effect models in multisite randomized trials. It presents various approaches to deriving parameter estimates by incorporating matrix and individual forms, balanced and unbalanced designs, different coding schemes, and some special cases. By showing the similarities, differences, and relationships among these estimates, the study provides a clearer framework, enhances transparency, and improves interpretability. This work can be extended to other designs in power analysis and optimal design, as well as broader research contexts. Future studies can further explore their applications across diverse study designs and practical settings.

**Acknowledgement**

Funding Statement: None

**References**

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>

Dong, N., Kelcey, B., & Spybrook, J. (2021). Design considerations in multisite randomized trials probing moderated treatment effects. Journal of Educational and Behavioral Statistics, 46(5), 527-559. <https://doi.org/10.3102/1076998620961492>

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). Applied longitudinal analysis. John Wiley & Sons. <https://DOI.org/10.1002/9781119513469>

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. Biometrika, 73(1), 43-56. <https://doi.org/10.1093/biomet/73.1.43>

Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. Biometrika, 74(4), 817-827. <https://doi.org/10.1093/biomet/74.4.817>

Mason, W. M., Wong, G. Y., & Entwisle, B. (1983). Contextual analysis through the multilevel linear model. Sociological methodology, 14, 72-103. <https://doi.org/10.2307/270903>

Raudenbush, S., & Bryk, A. S. (1986). A hierarchical model for studying school effects. Sociology of education, 59(1). 1-17. <https://doi.org/10.2307/2112482>

Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. Psychological methods, 5(2), 199-213. <https://doi.org/10.1037/1082-989X.5.2.199>

Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. Advanced Quantitative Techniques in the Social Sciences Series/SAGE. <https://psycnet.apa.org/record/1992-97494-000>

Snijders, T. A. B., & Bosker, R. J. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modeling (2nd ed.). Sage. <https://uk.sagepub.com/en-gb/eur/multilevel-analysis/book234191>