



Received: November 21, 2024

Revised: January 17, 2025 (1st); February 23, 2025 (2nd)

Accepted: February 23, 2025

Corresponding author:

Chiyeon Lim, Ph.D.

Stroke and Neurovascular Regulation Laboratory, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA
Tel: +1-617-407-9152

Fax: +1-617-407-4010

Email: CLIM8@mg.harvard.edu

ORCID: <https://orcid.org/0000-0003-0178-6976>

How to construct analysis of covariance in clinical trials: ANCOVA with one covariate in a completely randomized design structure

WooJin Jung¹, Kwan Lee², Hyung-Hwan Kim³, Chiyeon Lim³

¹Department of Biostatistics and Medical Informatics, Dongguk University School of Medicine,

²Department of Preventive Medicine, Dongguk University School of Medicine, Gyeongju, Korea,

³Stroke and Neurovascular Regulation Laboratory, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA

Analysis of covariance (ANCOVA) is a statistical method used to assess mean differences between groups by considering factors such as covariates or fixed effects and is often used to assess efficacy endpoints in clinical trials. When performing ANCOVA, the slope of the regression model should be the same for all treatment groups, with no interaction between the group and the covariate. Therefore, before analysis, the significance of the full ANCOVA model with interactions must be tested. If the interaction in the full model is statistically significant, the model that includes the interaction should be used; otherwise, ANCOVA using a reduced model without the interaction should be performed. If the ANCOVA model is not significant, this analysis method is not appropriate and a multivariate analysis or individual regression line estimation can be considered. If the difference in means between the groups is tested by ANCOVA, the confidence interval for the adjusted mean (least-squares mean) should be calculated and tested. Because the results may change depending on the covariates used in the ANCOVA model, the covariates should be predefined before performing the analysis. If a new covariate must be defined after a clinical trial is initiated, it should be specified in the statistical analysis plan. This is considered a major amendment; thus, the covariates must be redefined before clinical trial completion and must be described in the clinical study report. A clear report describing whether the redefinition of the covariates affected the sample size or decision-making is also necessary.

Keywords: Analysis of variance; Clinical decision making; Clinical trials; Covariates; Interaction; Linear model; Outcome assessments; Pretreatment expectations; Regression.



© The Korean Society of Anesthesiologists, 2025

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Analysis of covariance (ANCOVA) is often used to assess primary endpoints in clinical trials. The primary endpoint is a variable capable of providing the most clinically relevant and convincing evidence directly related to the major objective of a clinical trial and is often affected by various factors. ANCOVA is a representative statistical method used to assess mean differences between groups by considering factors such as covariates or fixed effects.

Assessing efficacy in clinical trials commonly involves comparing baseline to endpoint values through performing t-tests or analysis of variance (ANOVA) on mean changes. However, as these methods can be misleading with covariates, the use of ANCOVA should be considered. ANCOVA is a statistical method used to assess the effect of extra-

neous variables that cannot be controlled in experimental or correlation studies; however, it has a high risk of misuse [1].

When used properly, ANCOVA increases the power in clinical trials; however, its advantage over t-tests or ANOVA can be nullified when the sample size is small, assumptions are violated, or the analysis is used incorrectly [2]. Therefore, before analyzing with ANCOVA model, it should be checked carefully whether the covariate can be adjusted to evaluate the effect of the response variable is fundamentally reliable [3]. Statistical control can be obtained using ANCOVA to meet the unreliability of the pre-treatment values [4].

ANCOVA is achieved by regressing the post-treatment values for the pre-test values, which may be covariates within each treatment group. Therefore, the general assumptions of this test must be carefully considered [5,6]. In addition, with clinical trials designed to evaluate new drug applications, the statistical analysis methods should be transparent for the communication of the research results to be effective, as statistical analysis tests effectively communicate the study objectives [7]. As ANCOVA is a statistical analysis method that is adjusted using covariates, it can be used when the study outcomes are continuous, and logistic regression can be used when the outcomes are categorical in clinical trials. In this review, we address the limitations and requirements for using ANCOVA in confirmatory studies, such as Phase 3 randomized controlled trials.

Purpose of ANCOVA in clinical trials

As covariates are extraneous variables known to affect efficacy endpoints in clinical trials, they are difficult to control in an experimental design. Therefore, the effect of covariates on this efficacy endpoint within each treatment group is statistically controlled using a regression model and then tested between the treatment groups. The process for ANCOVA is the same regardless of the purpose of the analysis; however, the validity of the interpretation may differ. In general, ANCOVA is performed for the following three purposes: (1) To increase the statistical power of a

model by reducing the error term, (2) To adjust the heterogeneity of the comparison between treatment groups by statistically controlling covariates, and (3) To check the mean difference between the treatment groups, for which several prognostic factors (or extraneous variables) are set as covariates and used as means of statistical control.

Covariates

In clinical trials, in addition to the treatment group (independent variable) and response variable (treatment effect), extraneous variables can be measured and included in the model as covariates. Covariates are variables that can affect the response variable but are not affected by the treatment in clinical trials. The covariates used for adjustment should be specified in the clinical trial plan in advance based on clinical significance [8]. One reason for considering covariates is to reduce the error term in the ANCOVA model [9].

A within-group pretest-posttest design in clinical trials is easy to implement and consists of a pretest, intervention, and posttest. The pretest is conducted before the intervention is applied, which provides a baseline against which the subsequent test can be compared. Table 1 shows the within-group pretest-posttest design and the comparative between-group design.

As covariates should be measured before treatment, the baseline values of the study endpoint can be covariates and a reduction in errors can be expected when the response variable has a strong correlation. That is, the variables measured at baseline or pretest can be considered covariates in pretest-posttest clinical trials. In general, covariates used for ANCOVA can be considered when clinical significance is found in several studies.

ANCOVA model

ANCOVA is an extension of ANOVA for comparing mean differences among groups. From a methodological perspective it can be viewed as a combination of regression analysis and ANOVA.

Table 1. ANCOVA with Within-Group Pre-Posttest Design and Between-Group Comparative Design

Obs. (k)	Test group (i = 1)		Control group (i = 2)	
	Visit 1 (j)	Visit 2	Visit 1	Visit 2
1	y_{111}	y_{121}	y_{211}	y_{221}
2	y_{112}	y_{122}	y_{212}	y_{222}
3	y_{113}	y_{123}	y_{213}	y_{223}
⋮	⋮	⋮	⋮	⋮
n	y_{11n}	y_{12n}	y_{21n}	y_{22n}

i: group (test = 1, control = 2), j: visit (visit 1 = 1, visit 2 = 2), k: individual number (k = 1, ..., n). ANCOVA: analysis of covariance.

This review explains the simplest model using one independent variable, one response variable, and one covariate.

Model

The ANCOVA statistical model for a one-way layout with one covariate is calculated as follows:

$$Y_{ij} = \mu + \alpha_j + \beta_1 (X_{ij} - \bar{X}_{..}) + \varepsilon_{ij} \quad (1)$$

where Y_{ij} is the response variable of the i^{th} individual in the j^{th} group, μ is the overall true mean, α_j is the effect of group j , β_1 is the linear regression coefficient of Y on X , X_{ij} is the covariate for the i^{th} individual in the j^{th} group, $\bar{X}_{..}$ is the overall covariate mean, and ε_{ij} is the error associated with the i^{th} individual in the j^{th} group.

Assumptions

ANCOVA combines regression analysis and ANOVA, and the assumptions are equal or similar for both models. These include independence, normality, homogeneity of variance, linearity, and equal slope of the regression equation. An equal slope of the regression equation means that the slope for the response variable and covariates in all treatment groups is the same; therefore, no interactions are found between the covariates among the treatment groups.

Adjusted mean

An ANCOVA controlled for covariates is used to compare whether the difference in the mean of the response variables between the treatment groups is significant; therefore, the adjusted mean is used for testing. If no covariate effects are found, ANOVA is a more appropriate method. In addition, if the interaction is significant, the adjusted mean will be meaningless because the mean difference between the treatment groups varies according to the covariate. In other words, the interpretation of the adjusted mean is only meaningful when the slope of the regression equation does not differ within each treatment group.

When the ANCOVA model is the same as in Equation 1 above, the formula for the adjusted mean is as follows:

$$\text{adj } \bar{Y}_j = \bar{Y}_j - b_w (\bar{X}_j - \bar{X}_{..}) \quad (2)$$

where $\text{adj } \bar{Y}_j$ is the adjusted mean for the j^{th} group, $\bar{Y}_j$ is the unadjusted mean for the j^{th} group, b_w is the pooled within-group

regression coefficient, $\bar{X}_j$ is the covariate mean for the j^{th} group, and $\bar{X}_{..}$ is the overall covariate mean.

Hypothesis test for ANCOVA

Interaction

It should be checked whether there is an interaction between the covariate and the treatment group, which is an independent variable. For this, it is necessary to test whether the slopes between the treatment groups are the same:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_\alpha$$

If the interaction is not statistically significant, the assumption that the slope of each treatment group is the same is met. The significance of the full model, including the interactions, is tested. If the interaction is statistically significant in the full model, the model with the interaction should be used. Otherwise, ANCOVA should be performed using the reduced model without the interaction. Two types of comparisons may be performed when the slopes are not equal: one is to compare the distances among the regression lines for some covariate values, and the other is to compare the slopes or models estimated at the same covariate values [10].

Fig. 1 shows the relationship between the group means as a function of the covariate when the slopes are equal and unequal. When the slopes are unequal, the distance between the two regression lines depends on the covariate. This case is called a covariate-by-treatment interaction.

The slopes of the treatment groups must thus be tested to determine whether they are similar. If the interaction is not statistically significant, the assumption that the slope of each treatment group is the same is satisfied. The significance of the full model, including interactions, is then tested. If the interaction is statistically significant in the full model, the model with the interaction should be used. Otherwise, ANCOVA should be performed using the reduced model without the interaction.

Covariate test

Testing the effect of covariates on the reduced model (no interaction) involves determining whether the slope is zero in the regression equation between the response variable and the covariate within each treatment group:

$$H_0 : \beta = 0$$

If the slope of the regression equation is non-zero, a correlation is considered to exist between the response variable and the co-

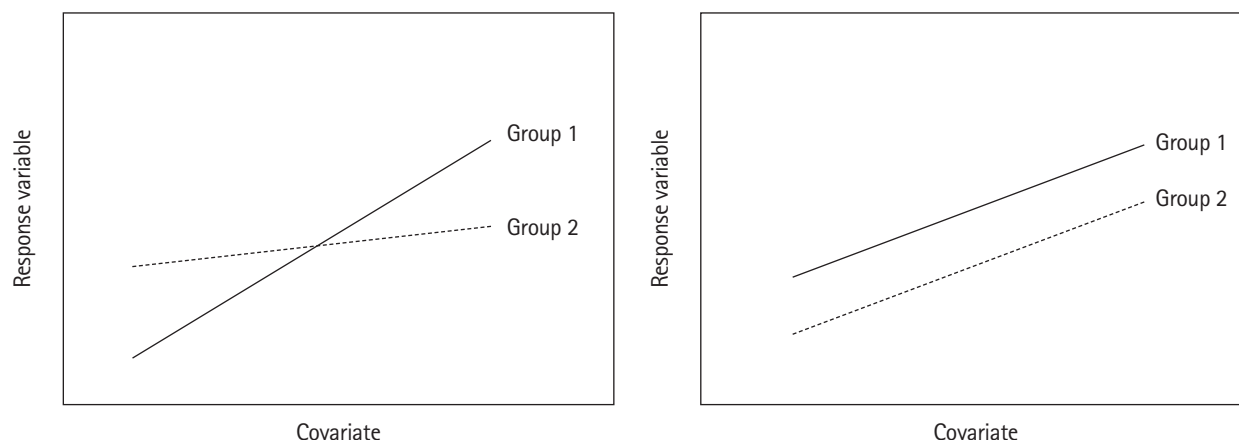


Fig. 1. Graphical representations of unequal and equal slopes.

variate. If the slope of the covariates in this test is zero, it would be more appropriate to perform an ANOVA than an ANCOVA.

Interaction model and unequal slopes model: covariate-by-treatment interaction

If the interaction between the treatment group and the covariate is significant, a model with the slope of the regression equation within each treatment group is considered more appropriate. In this case, the interaction model is applied and the adjusted mean is meaningless because the mean difference of each treatment group changes depending on the covariate. If the interaction is significant, the range of covariate values with significant differences between the treatment groups can be identified. Johnson and Neyman [11] and Potthoff [12] developed methods for establishing these domains; however, this is beyond the scope of this review and thus not covered.

Definition of the primary endpoint in ANCOVA

Mean change in the study endpoint

Confidence interval estimates could differ depending on the endpoint of interest (mainly the primary endpoint). Thus, before hypothesis testing or confidence interval calculations are considered, the predefined primary endpoint should be checked.

For example, suppose that the definition of the primary endpoint for comparison between groups is as follows: (1) comparing values from baseline at post-treatment ($Y_{\text{post-treatment}} - Y_{\text{baseline}}$), which may be defined as the within-group pre-post or post-pre; (2) comparing post-treatment ($Y_{\text{post-treatment}}$) with baseline (Y_{baseline}) values as a covariate; and (3) comparing values at post-treatment

from baseline ($Y_{\text{post-treatment}} - Y_{\text{baseline}}$) with the baseline values (Y_{baseline}) as a covariate, which may be defined as the within-group pre-post or post-pre.

For the first example above, ANOVA without covariance can be performed; however, if other covariates are present, ANCOVA is appropriate. For all three examples, the baseline value of the primary variable is often adjusted. However, because the statistical models are completely different, the three models cannot be considered identical, even though the estimates of the confidence interval for the true mean difference have the same values as those of the second and third examples above. As the models are different, proceeding with the same interpretation is not appropriate.

Difference between pre and post

The purpose of a randomized controlled trial is to conduct a hypothesis test on the differences between treatment groups. Therefore, the primary endpoint examples defined above are compared between groups, and the change within each treatment group is predefined by either 'pre minus post' or 'post minus pre.'

Another important consideration is the sign, as the absolute values may be the same but with the opposite sign. The direction should be accurately predefined in the protocol, and the values and characteristics should be reflected in the literature to determine clinically important thresholds, such as the non-inferiority margin.

Analysis

Statistical hypothesis testing for ANCOVA requires that the following assumptions be assessed: (1) when the P value of the entire model is less than the given significance level, the model is con-

Table 2. Analysis of Covariance Example Table

	P value from ANCOVA	LS Mean	SE	100 (1- α)% CI	
				Lower bound	Upper bound
ANCOVA parameter					
Group	x.xxxx				
Covariate	x.xxxx				
Least-squares means and group difference					
Test		x.xxxx	x.xxxx	x.xxxx	x.xxxx
Control		x.xxxx	x.xxxx	x.xxxx	x.xxxx
Test-Control	x.xxxx			x.xxxx	x.xxxx

ANCOVA: analysis of covariance, LS: least-squares, SE: standard error.

sidered significant; (2) the covariates provided must be independent of the treatment groups; and (3) the effect of covariates cannot be zero.

For the first assumption above, a post hoc test is not commonly performed. If the P value for the entire model is not significant (i.e., when conducting an ANOVA), testing the significance of the probability for the entire model is critical. For the second assumption, the value obtained after exposure to treatment is inappropriate as a covariate; thus, it is considered a pretreatment value in clinical trials. In the statistical model, this is tested through an interaction test; if the interaction is not significant, the ANCOVA model should present a model that analyzes only the main effect, excluding the interaction. Any covariates considered by ANCOVA may not be affected by the treatment [13]. For the third assumption, if the actual effect is zero, it does not need to be included in the model. If the second and third assumptions are not satisfied, ANCOVA is not appropriate and a multivariate analysis or individual regression line estimation can be considered.

Non-inferiority trial and ANCOVA

In order to show non-inferiority compared to a control, the lower or upper limit of a two-sided 100 (1 – α)% confidence interval for the true difference should be calculated. If the hypothesis to demonstrate non-inferiority is left-tailed, then the lower bound of the confidence interval must be greater than the non-inferiority margin, and if it is right-tailed, the entire upper bound must be smaller than the non-inferiority margin [14,15]. In this case, the confidence interval should be calculated using the statistical analysis method, and if the mean difference between groups is tested using ANCOVA, the confidence interval for the adjusted mean should be calculated and tested. In particular, comparing and interpreting the confidence interval and non-inferiority margin only when the definition of the study endpoint mentioned above is correct is acceptable. Table 2 presents an ANCOVA example table.

Conclusion

ANCOVA is used to adjust for differences in relevant baseline values between or among groups to improve not only the statistical power of significance tests, but also the precision of treatment effect estimates [13]. Thus, it is important to precisely show the adjusted treatment effect if an ANCOVA is performed for the primary objective. As the results may change depending on which covariates are used in the ANCOVA, the covariates must be defined in advance. If covariates need to be newly defined after a clinical trial has begun, it should be specifically described in the statistical analysis plan. As this corresponds to a major amendment, it should be reported to the related regulatory agency and the institutional review board. Moreover, the covariates must be redefined before a clinical trial is completed and a clear report describing whether the redefinition of the covariate affects the sample size and any decision-making should be provided.

Funding

None.

Conflicts of Interest

Chiyeon Lim has been an statistical editor for the Korean Journal of Anesthesiology since 2014. However, she was not involved in any process of review for this article, including peer reviewer selection, evaluation, or decision-making. There were no other potential conflicts of interest relevant to this article.

Data Availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Author Contributions

WooJin Jung (Conceptualization; Validation; Writing – original draft; Writing – review & editing)

Kwan Lee (Investigation; Writing – original draft; Writing – review & editing)

Hyung-Hwan Kim (Investigation; Writing – review & editing)

Chiyoon Lim (Conceptualization; Formal analysis; Investigation; Validation; Writing – original draft; Writing – review & editing)

ORCID

WooJin Jung, <https://orcid.org/0009-0003-3152-5344>

Kwan Lee, <https://orcid.org/0000-0003-2554-4170>

Hyung-Hwan Kim, <https://orcid.org/0000-0003-2980-0514>

Chiyoon Lim, <https://orcid.org/0000-0003-0178-6976>

References

1. Im S. Understanding and Application of Analysis of covariance (Research Methods 11). Seoul, Kyoyookbook. 2002.
2. Egger MJ, Coleman ML, Ward JR, Reading JC, Williams HJ. Uses and abuses of analysis of covariance in clinical trials. *Control Clin Trials* 1985; 6: 12-24.
3. Lamb J. Volume 3: Toxicology Testing and Evaluation. In: Comprehensive Toxicology. 2nd ed. Edited by McQueen C. Oxford, Elsevier. 2010, pp 1-2.
4. Sörbom D. An alternative to the methodology for analysis of covariance. *Psychometrika* 1978; 43: 381-96.
5. Reichardt CS, Mark MM. Nonequivalent group designs. In: International Encyclopedia of the Social & Behavioral Sciences. Edited by Smelser NJ, Baltes PB: Oxford, Pergamon. 2001, pp 10655-60.
6. Longford NT. Analysis of covariance. International Encyclopedia of Education. 3rd ed. Edited by Peterson P, Baker E, McGaw B: Amsterdam, Elsevier Science. 2010, pp 18-24.
7. Huitema BE. The analysis of covariance and alternatives: statistical methods for experiments, quasi-experiments, and single-case studies. 2nd ed. Hoboken, John Wiley & Sons, 2011; 1-618.
8. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH harmonised guideline: addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. ICH E9(R1) [Internet]. 2019 Nov 20 [cited 2024 Nov 21]. Available from https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf
9. Garson GD. GLM Univariate, ANOVA, and ANCOVA. Ashboro: Statistical Associates Publishing, 2012; 1-176.
10. Milliken GA, Johnson DE. Analysis of messy data, volume III: analysis of covariance. New York: Chapman & Hall/CRC, 2001; 1-624.
11. Johnson PO, Neyman J. Test of certain linear hypotheses and their application to some educational problems. *Stat Res Mem* 1936; 1: 57-93.
12. Potthoff RF. On the Johnson-Neyman technique and some extensions thereof. *Psychometrika* 1964; 29: 241-56.
13. U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Oncology Center of Excellence. Adjusting for covariates in randomized clinical trials for drugs and biological products: guidance for industry [Internet]. 2023 May [cited 2024 Nov 21]. Available from <https://www.fda.gov/media/148910/download>
14. Ko MJ, Lim CY. General considerations for sample size estimation in animal study. *Korean J Anesthesiol* 2021; 74: 23-9.
15. European Medicines Agency (EMA). Committee for Medicinal Products for Human Use (CHMP) [Internet]. Amsterdam: EMEA; January 1, 2020 [cited 2024 Nov 21]. Available from <https://www.ema.europa.eu/en/chmp>