Editorial

A Clinical-Statistical Case Study: Addressing Common Interpretive Ambiguities

Martino Recchia*

Medistat, Clinical Epidemiology and Biostatistic Unit, Milano; Mario Negri Institute Alumni Association (MNIAA), Italy

https://www.biolscigroup.us

Check for updates

Abstract

This article presents a hypothetical clinical scenario to illustrate interpretive ambiguities frequently encountered in mixed-design ANOVA. While based on a fictional dataset, the scenario serves to illustrate methodological challenges that arise when interaction effects are significant, but post-hoc tests yield non-significant results. The objective is to provide insight into this discrepancy and offer guidance for clinical researchers navigating similar situations.

Introduction

Statistical analyses, such as mixed-design ANOVA, are powerful tools for evaluating treatment efficacy in clinical trials. Mixed-design ANOVA is widely used in biomedical research for analyzing both within- and between-subject factors [1,2]. Yet, interpretive ambiguities often arise, especially when statistically significant interaction effects are not mirrored by post-hoc comparisons. These scenarios can puzzle researchers and potentially lead to misinterpretation. This editorial aims to clarify the rationale behind such inconsistencies, using a hypothetical example constructed specifically for this purpose.

Study description

The clinical scenario discussed here is entirely hypothetical and was conceived to illustrate a methodological issue. Consequently, no real-world patient data, treatment details, or ethical protocols are applicable since this is a fictional scenario.

Results

In our illustrative case, participants were assigned to three groups (placebo, treatment at 300 mg, and treatment at 450 mg) and evaluated at baseline and at the two-month follow-up using the SF-12 Quality of Life questionnaire [3].

To test the hypothesis of a potential dose–response relationship, a split-plot ANOVA (also known as mixed-design ANOVA) was employed, as it was deemed the most appropriate method for addressing the outlined analytical framework. The justification for using this design is as follows:

- Between-subjects factor: GROUP (Placebo, Dose1, Dose2)

- Within-subjects factor: TIME (Baseline, 2 months)

The Split-Plot ANOVA was selected as it enables (1) testing the main effects of both the Group factor and Time, and (2) examining the Time × Group interaction, which is the key component of the analysis.

Post-hoc comparisons: Tukey's test was used to determine differences between means. Tukey's HSD, one of the most conservative post-hoc tests, is less likely to detect smaller effect sizes [1,2], but only under specific conditions:

- When a significant interaction is present (to explore specific time points or groups)

- Or when a significant main effect is detected for any factor level

Advantages of this approach include:

001

- Properly accounting for the experimental design structure

- Correctly managing the longitudinal nature of the data (repeated measures)

- Providing statistically sound and clinically interpretable results

The statistical analysis revealed a significant interaction effect between Treatment and Time ($p = 0.0400$), indicating that treatment effects are time-dependent and that at least one specific combination differs meaningfully from the others.

In contrast, both main effects—Treatment and Time—were clearly non-significant ($p > 0.90$ and $p \approx 0.80$, respectively), indicating no overall difference when these factors were considered independently.

Subsequent Tukey's post-hoc comparisons, applied to explore the nature of the interaction, did not identify any specific pair of group-time combinations as statistically significant.

Statistical discussion

A significant interaction without significant post-hoc differences can occur when the global F-test captures subtle shifts spread across multiple conditions. Post-hoc tests like Tukey's test is conservative and may fail to detect subtle, distributed differences. The interaction F-test examines overall variance, while pairwise comparisons target specific group contrasts.

While an interaction plot is not included in this version, we suggest clinical researchers interpret such outcomes cautiously, integrating statistical and clinical insights, using clinical reasoning and alternative metrics like effect sizes or Bonferroni-adjusted LSD comparisons effect sizes or Bonferroni-adjusted LSD comparisons where appropriate [4-6], where appropriate.

Conclusion

This article highlights how statistically significant interactions in ANOVA can coexist with non-significant post-hoc results, emphasizing the need for careful interpretation. Researchers should:

1. Clearly report and interpret interaction effects.

2. Recognize the limitations of conservative post-hoc tests.

3. Use graphical representations and clinical context to enhance understanding.

Although hypothetical, this example mirrors real-world statistical challenges and underscores the value of integrating statistical insight with clinical judgment. These dilemmas are well-documented in literature, reinforcing the importance of precise statistical interpretation in biomedical research [1,2].

Disclaimer

No real patient data were used. The scenario is entirely hypothetical and intended for educational illustration only.

References

1. Field A. Discovering statistics using IBM SPSS statistics. 5th ed. London: SAGE Publications; 2018. https://www.scirp.org/reference/referencespapers?referenceid=3504991

2. Keppel G, Wickens TD. Design and analysis: a researcher's handbook. 4th ed. Upper Saddle River (NJ): Pearson Prentice Hall; 2004. Available from: https://www.proquest.com/docview/195085664?sourcetype=Scholarly%20Journals

3. Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34(3):220–33. Available from: https://doi.org/10.1097/00005650-199603000-00003

4. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310(6973):170. Available from: https://doi.org/10.1136/bmj.310.6973.170

5. Altman DG. Practical statistics for medical research. London: Chapman & Hall. 1991;285–288. Available from: https://www.scirp.org/reference/referencespapers?referenceid=1227285

6. Kirk RE. Experimental design: procedures for the behavioral sciences. Belmont (CA): Brooks/Cole Publishing Company. 1968. Available from: https://www.scirp.org/reference/referencespapers?referenceid=1238321