



Published in final edited form as:

J Paramed Sci. 2015 ; 6(3): 153–154.

Errors in statistical analysis and questionable randomization lead to unreliable conclusions

Brandon J George¹, Andrew W Brown^{1,2}, and David B Allison^{1,2,3}

¹Office of Energetics, University of Alabama at Birmingham, Birmingham, AL, USA

²Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, USA

³Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA

Dear Editor

We read with interest the paper, “The effect of food service system modifications on staff body mass index in an industrial organization” [1]. We noticed several substantial issues with data and calculations, calling into question the randomized nature of the study and validity of analyses. The distribution of baseline weight was significantly different between groups (p -value = “0.00”). We replicated the test using reported means and standard deviations (SDs) and obtained a p -value of approximately 1.9×10^{-17} . It is extraordinarily unlikely that any variable would be that different between two groups if allocation was truly random. Even if it was truly random, the stated method of “the samples were randomly divided into two groups” [1] does not describe the “method used to generate the random allocation sequence” and the “type of randomization; details of any restriction (such as blocking and block size)” details specified by Consolidated Standards of Reporting Trials (CONSORT) [2].

Given the large difference in baseline weights, it is unusual that the difference in baseline body mass index (BMI) between groups is not more significant ($p=0.032$), raising the question of what the groups’ distributions of height were. Both groups have 30 males (58.8%), so sex differences are unlikely to explain this discrepancy. Height was not explicitly reported, but it was possible to estimate height utilizing geometric means from body weight and BMI [3,4]. We calculated the baseline control group geometric mean as 2.04 cm taller than the test group. These calculations also suggest the control group shrunk by 1.26 cm while the test group grew by 1.52 cm over the study. Neither change is explained by rounding error nor seems plausible for adult subjects over 40 days.

Because there were no SDs of the change scores reported, we were unable to replicate the reported p -value (0.318) for the between-group test of weight change exactly. However, we were able to consider the pre and post-intervention SDs and calculate possible SDs of within-group change scores for a range of pre-post correlations. The largest p -value possible was 0.1282, calculated when each group had perfect negative pre-post correlation (correlation=-1), which is unlikely. If there was no or a positive correlation, the p -value

*Corresponding Author: brgeorge@uab.edu (B J George).

would be much smaller ($p=0.0449$ when correlation=0 for each group) and plausibly indicates a significant difference between groups. Therefore, although the published results are impossible the correct analysis could make the intervention appear more effective than reported.

The results section describes an initial sample size of 116 with 14 dropping out (p. 115). The tables report the remaining sample size to be 102, but the body of the text reports 101 subjects remained until study completion. It is unclear which values were correct; this lack of clarity also fails CONSORT guidelines [2].

Considering that the reported findings are essentially impossible given the stated study design, we encourage the authors to explain the treatment allocation and make the raw data available, or the journal to act according to the Committee on Publication Ethics [5] in situations where findings are unreliable.

Acknowledgments

Supported in part by NIH grants P30DK056336, R25DK099080, and R25HL124208. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or any other organization.

References

1. Hosseini R, Mirghotbi M, Pourvali K, Kimiagar SM, Rashidkhani B, Mirghotbi T. The effect of food service system modifications on staff body mass index in an industrial organization. *J Paramedical Sciences*. 2015; 6(1):113–118.
2. Schulz KF, Altman DG, Moher D. for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Ann Int Med*. 2010; 152 Epub 24 March.
3. Jean WH, Helms BP. Geometric Mean Approximations. *J Financial and Quantitative Analysis*. 1983; 18(3):287–293.
4. Wuyts, M. Mean of Ratios or Ratios of Means?. *Thinking with Data*. 2000. Retrieved from <http://development.thinkingwithdata.com/show.php?p=note.html>
5. Committee on Publication Ethics. Retraction Guidelines. 2009. Retrieved from <http://publicationethics.org/files/retraction%20guidelines.pdf>