



What is Compared in a Clinical Trial?

Emin Mehmet Yusuf*

Istanbul Universitesi – Cerrahpasa, Lisansüstü Eğitim Enstitüsü, Turkey

***Corresponding author:** Emin Mehmet Yusuf, Istanbul Universitesi–Cerrahpasa, Lisansüstü Eğitim Enstitüsü, Deniz Sk. Oymak Sitesi, No. 5, 34464, Yeniköy, Istanbul, Turkey, Tel: +90-212-292 6161; Fax: +90-212-292 8810; Email: mehmet.yusuf@altis.org

Short Communication

Volume 3 Issue 1

Received Date: February 05, 2020

Published Date: March 04, 2020

DOI: 10.23880/aabsc-16000145

Abstract

In a randomized clinical trial (RCT), there are universally accepted rules of thumb for choice of alpha (α : 0.05) and power ($1-\beta$: 0.80). These choices require consideration and need to be anything but automatic. Next level of consideration should be given to what we actually chose to compare: the value / effect of the intervention per se or that of the intervention strategy. There is a subtle difference which hinges on which dataset we analyze. Possible data sets are intention to treat (ITT), per protocol or on treatment (PP), as treated (AT). And on the highest level, are we interested in how likely we are to observe the data (that suggests the intervention has some effect) in the event that the null hypothesis (that our intervention has no effect) is, in fact, true or as an entirely different construct, how likely is it that the intervention has some effect given our observed data?

These are fundamental questions that we appear to decide on, when we chose alpha (α : 0.05) and power ($1-\beta$: 0.80) automatically, without giving them the thought process they deserve.

Keywords: Randomized; Trial; Meaning; Frequentist; Bayesian; Power

Decision on Alpha (α) and Power ($1-\beta$)

In a randomized clinical trial (RCT), alpha (α) and power ($1-\beta$) chosen by the investigator set an entire model of decision much like our judicial system. In the currently accepted Western judicial system, we do not want to find a defendant guilty when in fact that person may be innocent, and to that end, we may agree that the system may let some guilty defendants go free. In terms of clinical trial decisions, we do not want to make a type I error. Type I error is when we decide that the evidence is overwhelming when in fact it is not. The way to prevent this is to set alpha very low (not decide against the innocent unless evidence is beyond doubt). The rest of the analogy is lax power (letting some guilty get away), agree to a higher possibility of type II error.

In clinical trials we have the luxury of deciding on our “judicial system” case by case. Especially for those outcomes where the possible benefit is large, risk and cost are small, we should reverse our judgment system so as not to miss any

“guilty” party and agree to jail some “innocent”, meaning use a high alpha like 0.1 – 0.2 and a high power like 0.95 – 0.99. This is worth thinking about.

Choice on which Dataset to Compare

The data set used in the analysis, as well as the analytical method may also be decisive agents, in the investigator’s quest to reach a realistic estimate. Here lies another fundamental question: Do we want to compare the effect of the intervention against the control (inherent in the general definition of clinical trials) or the strategy of intervention against control. To clarify, if there is a treatment protocol, this is a strategy. The treatment may fail, be replaced by another treatment, the patient may fail to comply etc. None of this changes the physician’s initial treatment strategy. If we want to compare a treatment strategy against a control strategy than the dataset to analyze is what is known as intention-to-treat (ITT) [1,2]. This dataset has statistical advantages too in terms of bias.

On the other hand, if we want to compare the effect of the intervention against the control, then we cannot have treatment reversals, missing data etc. in our dataset, we need to analyze the per protocol (PP) dataset [3]. PP dataset only includes volunteers who adhered to the protocol for the entire duration of the trial. PP dataset thus removes noise from the signal, but as some point out, maybe, at the cost of possible introduction of bias. However, if we want to know how effective the intervention would be, if followed as instructed, that may be the best choice of datasets to analyze.

A third possibility is to analyze the final observation dataset, where everyone ends up eventually, the as treated (AT) dataset. This dataset, unlike the PP does not leave data unanalyzed. All volunteers are analyzed in the final group they complete the study.

These may, and at given conditions, will produce different results. This is worth thinking about.

Data Given Model or Model Given Data – Frequentist or Bayesian Design and Analysis

Another important decision is the information we chose to seek. Is it the level of improbability of the observed data (suggesting effectiveness of the intervention) if the intervention has no effect (low alpha / p value) or the likelihood of the effectiveness of the intervention based on the observed data (maximum posterior probability)?

The usual method we use to test a hypothesis employs a model that is not entirely intuitive. We declare a cutoff point (alpha). If our observations appear to fall within the confines of a predefined model (the null hypothesis – that suggests the intervention has no effect) by less than that cutoff point we conclude that given the null model, our observations cannot be, therefore the null model is not valid for our observations. Data given the model is not valid; therefore another model must be valid.

Our intuition generally works the other way around, we usually think, given this observed data, what is the most likely model? This, in a nutshell is the fundamental difference between frequentists and Bayesian approaches. There are many well established biostatistical methods and analytic tools that work with Bayesian statistics [4,5].

Is Conservatism Good – Should it be the Norm in Biostatistics?

Conservatism is generally well accepted in statistical analysis. Most authors would prefer an understatement to an exaggeration. In that spirit the ITT dataset is praised.

As it does include dropouts, omissions, execution errors all in the analysis, the final result is expected to be diluted and to a degree conservative. But is this always favorable? In superiority trials, where the objective is to show the new intervention is better than the old in some aspects, maybe. But in the past decade or two, as clinical trials evolved and became so successful, superiority trials became exceedingly difficult as margins grew ever narrower. Another type of clinical trial started to become more and more prominent: non-inferiority trials. In a non-inferiority trial, the objective is to show that the new intervention is not worse than the old intervention by a predefined margin (delta, e.g. % 20). The design and interpretation of this type of study is different. In these studies, if you fail to show the difference between the two groups (by more than delta) then you are successful! In the perspective of these studies, ITT dataset is the aggressive not conservative dataset to analyze. AT appears best suited to this scenario. Again, we need to think about what it is we are comparing before deciding on the dataset to analyze or the method to use [6].

Conclusion

When we accept the conventional alpha (α : 0.05) and power ($1-\beta$: 0.80) we appear to have chosen less type I error than type II error and the frequentist (data given model) approaches and possibly the ITT dataset. A great number of good books suggest these as the “standard” choices to make. The author believes, these choices should not be arbitrary and of course not standard, and should receive the thought process they duly deserve.

References

1. Wang Duolao, Bakhai, Ameet (2006) Clinical Trials, A Practical Guide to Design, Analysis, and Reporting, pp: 83-84
2. Piantadosi Steven (2017) Clinical trials: a methodologic perspective, 2nd (Edn.), pp: 402.
3. Chin Richard, Lee Y Bruce (2008) Principles and practice of clinical trial medicine, Academic Press, pp: 323
4. Stone James V (2013) Bayes' Rule: A Tutorial Introduction to Bayesian Analysis, Computer Science.
5. Navarro Danielle (2014) Foxcroft, David. Learning statistics with jamovi: a tutorial for psychology students and other beginners.
6. Friedman LM, Furberg CD, DeMets D, Reboussin DM, Granger CB (2015) Fundamentals of Clinical Trials, 5th (Edn.), Engineering & Materials Science, pp: 300.