Study Development Suggestions for CYP-Guides Clinical Trial Using: Power Analysis

This is a continued analysis from my other [.Rmd document](https://github.com/kittysocks/portfolio/blob/main/R/CYP-GUIDES-Power-Analysis.docx).This is still a **work in progress**. Data description will be repeated below for convenience:

# Data Description:

Source: [Kaggle](https://www.kaggle.com/datasets/shashwatwork/clinical-dataset-of-the-cypguides-trial/data)

### Context

CYP-GUIDES (Cytochrome Psychotropic Genotyping Under Investigation for Decision Support) is a randomized controlled trial (RCT) comparing 2 outcomes in hospitalized patients with severe depressive disorders treated according to the patient’s CYP2D6 genotype and functional status versus standard psychotropic therapy. The primary outcome was hospital Length of Stay (LOS) and the secondary outcome was the Re-Admission Rate (RAR) 30 days after discharge.

The trial setting was the Institute of Living at Hartford Hospital. CYP2D6 genotyping was implemented to characterize the functional status of the CYP2D6 enzyme with sub-normal, normal, or supra-normal function. The electronic medical record (EMR) was utilized to transmit clinically actionable drug prescribing guidance based on the patient’s CYP2D6 function to the physician.

### Content

The RCT recruited 1500 patients, genotyped CYP2D6 in 1459, and randomized 477 to standard therapy (Group S), for whom treatment-as-usual guidance was delivered without consideration of patient CYP2D6 genotype, and 982 to genetically-guided therapy (Group G) where CYP2D6-based treatment recommendations were provided via EMR to physicians. For inpatients in Group G whose CYP2D6 function was sub- or supra-normal, medications primarily metabolized by the CYP2D6 enzyme were proscribed.

The RCT developed a database of potential benefits to the field. The pharmacologic, clinical course, and pharmacogenetic therapeutic guidance is being published in a related article. These data should enable various investigators to assess effeeffects of clinical decision support on resource utilization and psychotropic therapy during psychiatric hospitalizations.

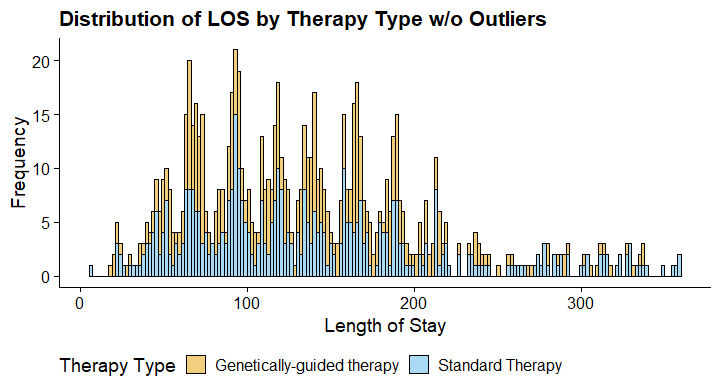
### Acknowledgements

Tortora, Joseph; Robinson, Saskia; Baker, Seth; Ruaño, Gualberto (2020), “Clinical Dataset of the CYP-GUIDES Trial”, Mendeley Data, V1, doi: 10.17632/25yjwbphn4.1 Source

## Distribution of Data

I am currently analyzing if the type of therapy has an effect on LOS of patients.

Below is the distribution of LOS among the two types of therapy, standard therapy and genetically-guided therapy. For graphing purposes, the genetically-guided therapy group was stratified and randomly sampled to match the sample size of the standard therapy group. Each therapy type has 477 individuals.



The outliers of the data were also removed for clarity. There were 43 outliers in the genetically-guided therapy group and 38 outliers in the standard therapy group.

We can see that there is a large overlap of the two distributions, however we can see small differences in the distributions. Our hypothesis for the LOS is below:

Null Hypothesis: There is no difference between the length of stay between the two therapies.

Alternative Hypothesis: There is a difference between the length of stay between two therapies.

## T-Test

From my other .Rmd, I conducted a T-Test between the two therapy groups, but I didn’t consider larger sample size of the genetically-guided therapy group. I could have performed a Welch’s T-Test, but instead I re-ran the test for the newly randomly sampled data.

t.test <- t.test(df$LOS ~ df$Assignment)  
p\_value <- t.test$p.value

Compared to doing a T-Test on the entire data (“0.515656957537065”), we got a p-value of 0.9412888. This p-value is not statistically significant.

If the p-value is not statistically significant, it means that, under the assumption that there is no difference in LOS between the two types of therapy, the observed data would occur at least 94.1288779% of the time. Therefore, the observed difference is not considered statistically significant. However, how can we know that this p-value was found in good faith? Did we have enough data to make a good decision? It’s important to consider whether the non-significant p-value reflects a true absence of effect or if it could be due to limitations in the study design, sample size, or other factors.

## Power Analysis

Power is the probability that a statistical test will detect a true effect or difference when it exists. In our context, it’s the probability that a statistical test will show significant difference between the LOS of the two therapy groups.

A Power Analysis will determine what sample size will ensure a high probability that we can be given results in good faith for the next time this experiment is performed. If we use the sample size recommended by the Power Analysis, we will know that regardless of the p-value, we can know that we used enough data to make a good decision.

In order to do a Power Analysis, we want to know how much Power we want. A common value is 0.8. That means we want an 80% probability that we will see statistical significance. If we ran the experiment 100 times, 80 of the experiments would show statistical significance. After determining power, we need to determine the threshold for significance, which is called alpha. A common value for alpha is 0.05. We then need to determine the effect size, which is the difference between the two observation’s means and standard deviation. The effect size is the overlap of data between the two groups. We can determine effect size by using the below formula:

Effect size:

Pooled estimate of standard deviations can be calculated using this:

There are a lot of different ways to find effect size. It depends on your data. Sometimes through literature, you might have to make an educated guess. But luckily, we have the mean and the standard deviations of the two groups from the experiment.

Table 1. Descriptive statistics of the CYP-GUIDES study

| **Therapy Type** | **Mean LOS** | **Standard Deviation** |
| --- | --- | --- |
| Genetically-guided therapy | 173.4193 | 188.4277 |
| Standard Therapy | 172.6080 | 149.4878 |

Pooled estimated standard deviations:

Effect Size:

The effect size found is super small. We could technically use Cohen’s “small” effect size, but I’m unsure of the complications of that.

So, now that we have the power, alpha, and effect size, we can determine power. We will be using functions from the “pwr” library in R.

##   
## Two-sample t test power calculation   
##   
## n = 741859.2  
## d = 0.0046  
## sig.level = 0.05  
## power = 0.8  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

From the Power Analysis, we can see that we would need a sample size of **74,1859** individuals per therapy type group to have a 80% probability of seeing statistical differences between the LOS in the two therapy groups. This is an impossible feat for this study, as the cost, and data integrity of 74,1859 patients seems very difficult to maintain. This sample size suggestion is not feasible.