

Power Analysis Consulting Report

MSSP Consulting: Zihuan Qiao, Anna Cook, Yinfeng Zhou, Zixuan Liu, Jiaheng Li

3/1/2021

Project Description

The client is interested in determining the relationship between nonsteroidal anti-inflammatory drugs (NSAIDs) and pain/disability measures in patients with radicular back pain, or pain that radiates from the back through the legs. This type of pain is often due to spinal stenosis and is typically treated with a combination of Gabapentin and physical therapy. Previous research suggests that NSAIDs may provide patients with additional relief, but this evidence is weak. The client's research aims to explore this relationship further with a sufficiently highly powered study.

The client's experiment will consist of two groups of patients with radicular back pain: one treatment and one control group. Both groups will be treated with a combination of Gabapentin and physical therapy. In addition, the treatment group will receive a regimen of Diclofenac (one type of NSAID), while the control group will receive an equal regimen of a placebo. Patients' outcomes will be measured through two different self-report scales. The first is a Visual Analog Scale (VAS) in which patient's report their pain level on a 0-10 scale. The second is the Roland Morris Disability Questionnaire, which is a series of 24 yes/no questions indicating the severity of a patients' disability. This measure will be scored as a sum of all of the "yes" responses. Scores for each of these two measures will be collected at the start of the treatment regimen, after one week, and after one month. According to previous research, a "clinically meaningful improvement" in scores is 30% or more.

The client's question for our team is how to determine a sufficient sample size such that it is possible to detect a 30% difference in scores between the treatment and control groups over the treatment period. To this end, we conducted a series of power analyses, described below.

Power Analysis

In order to determine a sufficient sample size for the client's experiment, we conducted a series of one-way, two-sample power t-tests. A two-sample test was chosen because the main comparison being made is between the control group and the treatment group which are two separate, independent samples. Additionally, the test must be one-way because the client is only interested in whether the treatment group shows an improvement in outcomes compared with the control group, so the t-test is one-sided.

The first power analysis was conducted using an alpha level of 0.05, effect size of 30%, and 80% power. A minimum of 80% power is a widely agreed-upon standard in the statistical literature, and alpha level of 0.05 is a commonly used threshold for establishing statistical significance. The results of this analysis are shown in Figure 1. Based on this analysis, the minimum optimal sample size is 139 subjects. It is important to note that this sample size is for each group and must be doubled to determine the total participants needed.

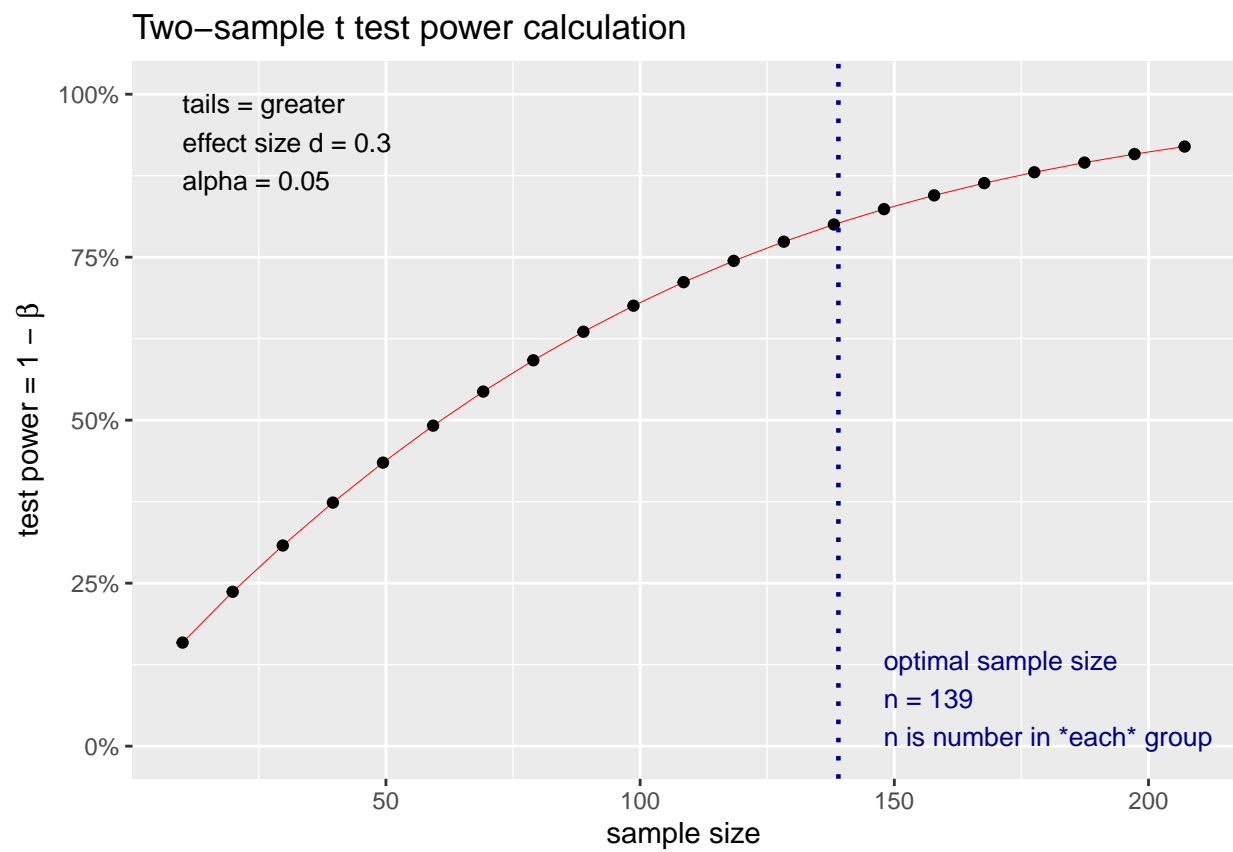


Figure 1: Power analysis for 30% effect size and alpha level 0.05 yields an optimal sample size of 139.

Next, we conducted a similar power analysis as above, this time using an alpha level of 0.01 which is another common threshold for establishing statistical significance. The results are shown in Figure 2. From this, we can see that the optimal sample size is 225 subjects per group.

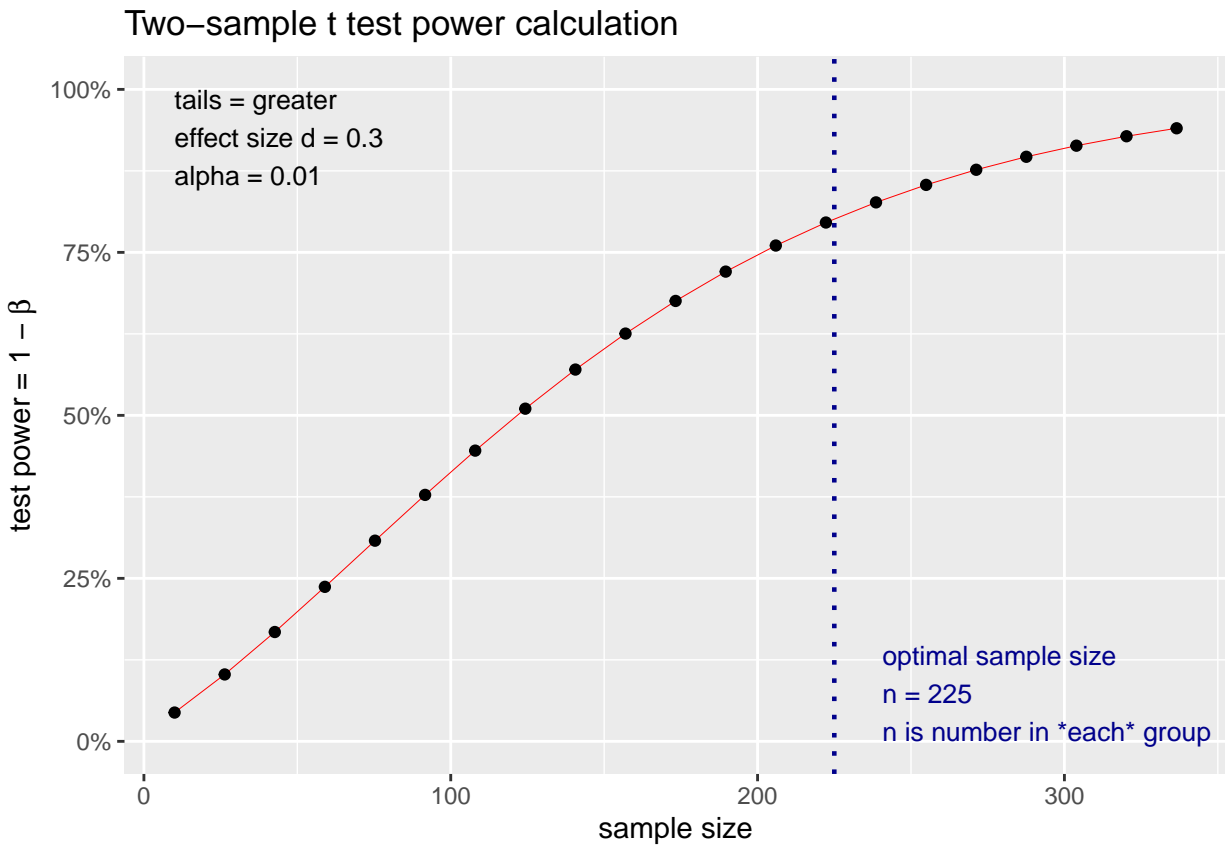


Figure 2: Power analysis for 30% effect size and alpha level 0.01 yields an optimal sample size of 225.

In addition to the power analyses conducted for 30% effect size, we repeated the same procedure for 50% effect size, per the client's request. Like the previous two analyses, we used 80% power and alpha levels of 0.05 and 0.01. The results are shown in Figures 3 and 4 respectively. The results of these power analyses show that the minimum necessary sample size to detect a 50% effect size is smaller than for a 30% effect size. This makes sense since a larger difference in scores between the treatment and control groups will be easier to reach statistical significance than a smaller difference in scores.

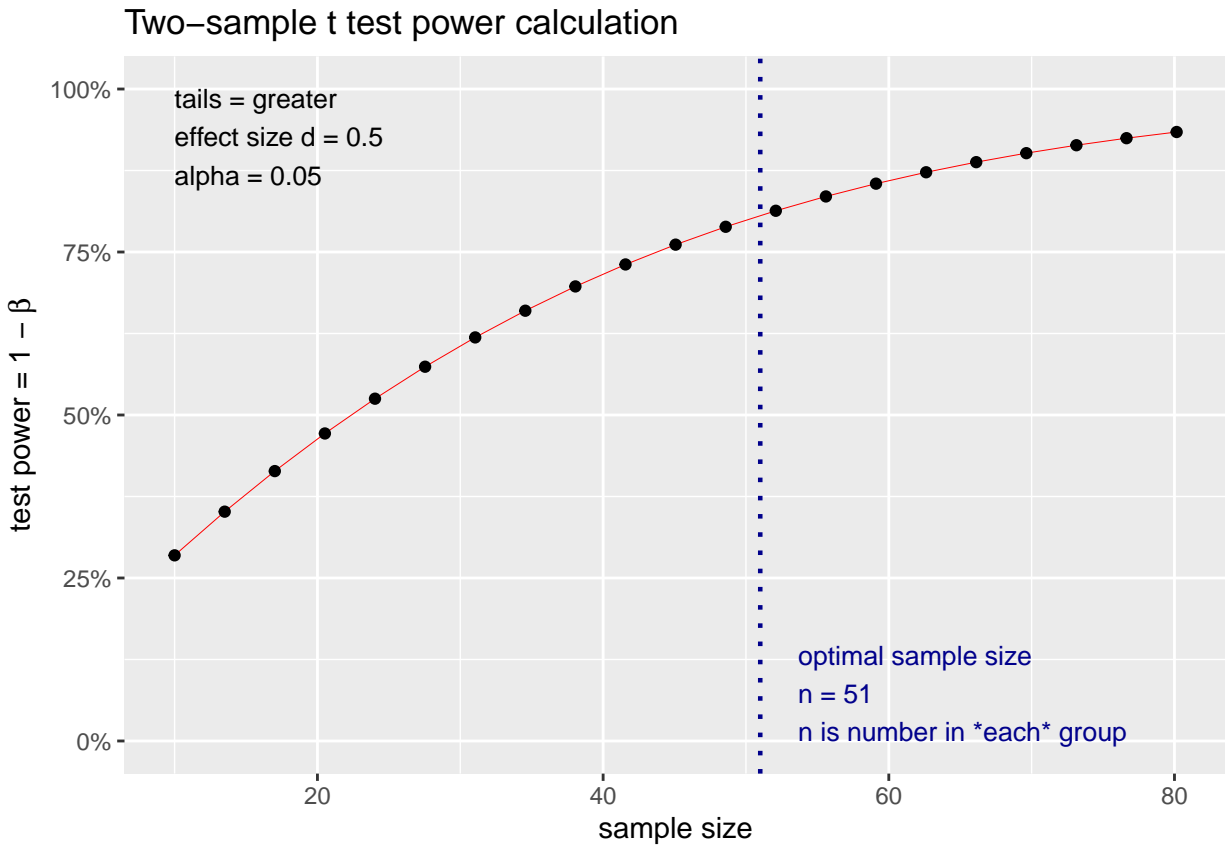


Figure 3: Power analysis for 50% effect size and alpha level 0.05 yields an optimal sample size of 51.

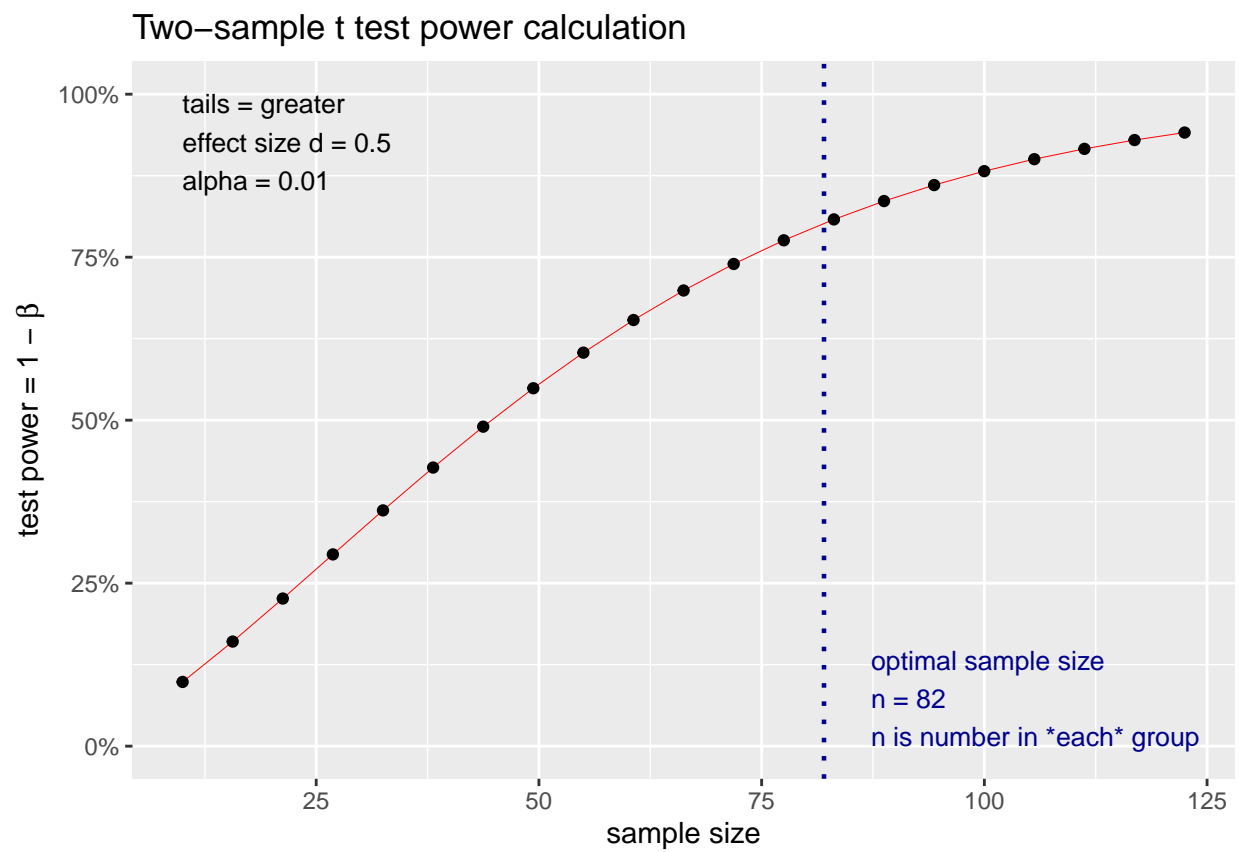


Figure 4: Power analysis for 50% effect size and alpha level 0.01 yields an optimal sample size of 82.

Conclusion

In conclusion, in order to achieve 80% power for a 30% effect size at alpha level of 0.05, assuming all assumptions hold, the minimum sample size required for the client's study is 139 participants per group, or 278 participants total.

It is important to note that the client has two different outcome measure that may be used in the study. Because the "clinically meaningful improvement" of 30% is common between the two measures, there is no difference in the power analyses for these two scores. The same number of participants will be required regardless of the client's choice of disability index.

References

Effect size and MCID:

- https://journals.lww.com/ejanaesthesiology/fulltext/2016/01000/the_minimum_clinically_important_difference_is.12.aspx
- <https://academic.oup.com/jpepsy/article/34/9/917/939415>

Client's background research:

- https://journals.lww.com/spinejournal/Fulltext/2008/01010/Minimal_Clinically_Important_Change_for_Pain.15.aspx?casa_token=DaFxxS8YJP8AAAAA:P5_K2_8G8Y6j26_an8kQ5eeGYUuMfC-Gqd5uUjMiZdpvTeZkr6NVeo_gwkfS-0E_eQNBxfixCJdLl274CiJsDPV8
- <https://www.sciencedirect.com/science/article/abs/pii/S1526590007008991>
- <https://www.nejm.org/doi/10.1056/NEJMoa1313265>