Global Journal of Medicine & Public Health

Sample Size Estimation for EORTC QLQ-C30 Summary Score as Primary Endpoint

Rajneesh Singh[1], Shalini Chandra[2]

ABSTRACT

Background

This paper discusses the importance of sample size in research, highlighting the ICH Guidance's requirement for clinical studies to clearly explain their sample size, particularly when using the EORTC QLQ-C30 Summary Score as the primary outcome measure. The idea behind this paper is to give the sample size when the outcome is Global Health Status and Out-of-Pocket Expenditure.

Method

The sample size calculation is based on a simulation technique using SAS software, assuming equal allocation between groups to yield a significant result with sufficient power.

Results

Sample sizes obtained using a simulation-based method require 25 patients per group to account for a minimum clinically significant difference of 14 units between the two groups, with 88% statistical power at a 5% level of significance.

Conclusion

The simulation-based technique, combined with validated software, can be helpful when the effect size and variability of the QoL score are not precisely known, particularly when the effect size is not sufficiently specified to determine the sample size accurately. Compared to conventional techniques, simulation-based sample size estimation increases confidence in reaching statistical power, particularly for complicated endpoints like the EORTC QLQ-C30 summary score.

Keywords: EORTC, QoL, Simulation, SAS, Sample Size.

GJMEDPH 2025; Vol. 14, issue 5 | OPEN ACCESS

1*Corresponding author: Rajneesh Singh ,Research Scholar- Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan, Email id- Rajneesh.biosts@gmail.com; 2. Shalini Chandra, Professor and Head, Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan

Conflict of Interest—none | Funding—none

INTRODUCTION

Researchers should evaluate the sample size during the planning and design phase of a clinical study to ensure the accurate power and precision of clinical research [1]. A sample size is crucial for researchers to evaluate the project's feasibility, cost, and time. Sometimes, it is complex when researchers do not have prior information regarding the component of sample size [1, 2, 3]. This paper presents a general approach for determining the sample size for health-related quality of life measures used in cancer patients. The 30-item EORTC QLQ-C30 is a disease-specific tool used to evaluate many aspects of cancer patients' quality of life. The Quality of Life score is essential in any clinical study, and it is a meaningful endpoint that should be able to accurately and consistently evaluate a patient's functions or survival. The patient-reported outcome, such as the EORTC-QLQ 30 Questionnaire, is valid for assessing patient experiences, functioning, or survival in a precise, accurate, and consistent manner [4, 5]. The EORTC QLQ-C30 instrument has been endorsed by the US FDA's Core Patient-Reported Outcomes in Cancer Clinical Trials Guidance for Industry [5, 6]. Five different functioning scales measure physical, role, emotional, cognitive, and social functioning [6]. Three symptom measures and quantify pain, nausea/vomiting, and exhaustion. Six single items on the questionnaire evaluate issues associated with cancer, such as dyspnea, difficulty sleeping, appetite loss, constipation, diarrhea, and financial troubles; they can be found in detail in Table 01 [4].

Table No.1: The EORTC QLQ-C30 Manual Score [4, 5, 6]

		Scale	No. of Question	Range	Question Number
Global health status / QoL	Global health status/QoL (revised)	QL2	2	6	29 & 30
Functional scales	Physical functioning (revised)	PF2	5	3	1,2,3,4,5
	Role functioning (revised)	RF2	2	3	6,7
	Emotional functioning	EF	4	3	21,22,23,24
	Cognitive functioning	CF	2	3	20 & 25
	Social functioning	SF	2	3	26 & 27
Symptom scales/items	Fatigue	FA	3	3	10,12,&18
	Nausea and vomiting	NV	2	3	14 &15
	Pain	PA	2	3	9&19
	Dyspnea	DY	1	3	8
	Insomnia	SL	1	3	11
	Appetite loss	AP	1	3	13
	Constipation	CO	1	3	16
	Diarrhea	DI	1	3	17
	Financial difficulties	FI	1	3	28

The EORTC QLQ-C30 scoring manual will be used to score data, transforming scales and items into 0-100 scales. Total scores will be calculated from categorical scales, including global health status, physical, role, emotional, cognitive, and social functioning. Symptom scales include Fatigue, nausea, vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties. Using the assessed qualitative scales, the following total scores will be determined for the global health-related quality of scale, the functional scales, and each symptom scale or item:

Table #2: The EORTC QLQ-C30 Scoring Procedures [4,6]:

Sr. No.	Scale	Subscale	Formula
1	**Global health status:**	*Global health status/QoL*	((Q29+Q30)/2-1)/6*100
2	**Functional scales:**	*Physical functioning*	(1-((Q1+Q2+Q3+Q4+Q5)/5-1)/3) *100
		Role functioning:	(1-((Q6+Q7)/2-1)/3) *100
		Emotional functioning	(1-((Q21+Q22+Q23+Q24)/4-1)/3) *100
		Cognitive functioning	(1-((Q20+Q25)/2-1)/3) *100
3	**Symptom scales/items:**	*Fatigue*	((Q10+Q12+Q18)/3-1)/3*100
		Nausea and vomiting:	((Q14+Q15)/2-1)/3*100
		Pain	((Q9+Q19)/2-1)/3*100
		Dyspnea	((Q8-1)/3*100
		Insomnia	(Q11-1)/3*100
		Appetite loss	(Q13-1)/3*100
		Constipation	(Q16-1)/3*100
		Diarrhea	(Q17-1)/3*100
4	**Financial Difficulties**	*Financial difficulties*	(Q28-1)/3*100

2.0 Conceptual Framework: Several factors are considered in a sample size calculation, including the study's objective, design, data analysis method, Type I error, Type II error, variability, and effect size [6, 7]. The study elements used are listed in the Table below, which can help understand the calculation of the sample size of the study.

Table No. 3: Prior Information Required for Sample Size Estimation [7, 8]

Pre-Request for Sample Size Calculation	Description through an example:
What is the primary purpose or aim of the study	To demonstrate the Global health status/QoL score in Cancer Patients [Example: Testing of Mean/ Proportion]
What is the primary measurement utilized to evaluate patient outcomes	Example: Global Quality Index from EORTC 30 Score
How will the information be analyzed to find any variations between the groups?	Example: T-statistic / or other statistics will be used to compare the between-groups for the Global health status/QoL score
What types of results may one anticipate?	Example: No Difference in Global health status/QoL score between the groups
The importance of minimal clinical difference	Example: 14 units + positive side or Minimum significant difference based on clinical inputs.

Before determining the sample size, a few preliminary steps, such as establishing the hypothesis and checking for statistical errors (Types I & II) should be defined. The statistical test's hypothesis setting is explained in the Table below.

Table # 4: Hypothesis Setting and Errors [9, 10, 11, 12]:

Research	Hypothesis Setting	Real Case	
		Ho is True	H1: is True
	Ho: Is True	Accurate 1- α	False Negative Type II Error β
	H1: Is True	False Positive Type I Error α	Accurate 1-β

3.0 Prior Information and Effect Size [13, 14, 15]: Researchers often possess extensive expertise in the same subject, which they can utilize to determine the effect size for the studied sample. Therefore, knowledge and research experience play a crucial role in adding to the subjective aspect of the experiment. There are several ideas regarding the effect size in practical contexts. One such strategy is determining the effect size corresponding to the minimal clinical significance. Because this minimal effect size is so small, finding it often requires considerable effort. An alternative approach is to evaluate the true magnitude of the underlying influence objectively. Table 5 presents six practical strategies that are suitable for use in practice, but are not comprehensive in their application to choosing the effect size.

Table No. 5: Practical Approaches to Evaluate the Effect Size Factors

Type of Evaluation	Which question should a researcher ask?
The smallest effect size of interest	How small an effect size is deemed notable from a conceptual or practical point of view?
The minimal statistically detectable effect	What critical effect size can be statistically significant given the test and sample size?
Expected effect size	According to prior studies or theoretical projections, what magnitude of effect is anticipated?
Width of the confidence interval	Which effect sizes are excluded according to the anticipated width of the confidence range surrounding the effect size?
Sensitivity power analysis	When performing a hypothesis test, which effects does a design have enough power to detect over a range of potential effect sizes?
Distribution of effect sizes in a research area	In a particular field of study, when effects are a priori unlikely to be found, what is the empirical range of effect sizes?

4.0 Statistical Method [10, 11]: Assume that we have two independent sample sizes of QoL data, each of size n for two groups like X and Y i.e. $x1, x2, \ldots, xn, and\ y1, y2, \ldots, yn.$
Global QoL summary score data that are continuous variable with cumulative distribution functions (cdfs)

are $Fx\ and\ Fy$ respectively.

Assuming cases when the distributions may differ in location but have a similar form and the δ is the location of the difference $i.e., average\ (y) - average\ (x) = \delta$

The major objective of the study will be to test the δ: $HA: \delta \neq 0$, with the null hypothesis, $H0: \delta = 0$.

A suitable test function (such as the t-test) to examine these hypotheses when comparing two groups to their anticipated mean score.

Statistical Formula for the Sample Size Estimation [12, 13, 14, 15, 26]:

The sample size required to reject the null hypothesis under specific assumptions is determined using the calculations below. For example, the formula is applied when comparing the means of two groups in an experiment. Suppose,

$\mu 0: Mean\ of\ Global\ Health\ Status\ First\ Group,$

$\mu 1: Mean\ of\ Global\ Health\ Status\ Second\ Group,$

$\sigma 0 = Standard\ Deviation\ of\ Global\ Health\ Status\ First\ Group,$

$\sigma 1 = Standard\ deviation\ of\ Global\ Health\ Status\ the\ second\ group$

$$N = 2 * CV^2 \left((t\alpha, n-2) + + (t\beta, n-2)\right)^2 / \Delta^2$$

$$Where\ \Delta = 0.2, CV = \frac{\sigma}{\mu R}$$

$For\ Two\ Sample\ Tests, 90\ \%\ CI\ should\ contain\ a\ 20\ \%\ difference.$

5.0 Results: This sample size calculation is based on a two-sided t-test statistic, based on the pilot study results where we assessed the global health status summary from EORTC- QLQ 30 Questionnaire. Calculated the sample size based on different effect sizes in Tables 6 and Figure #1 below.

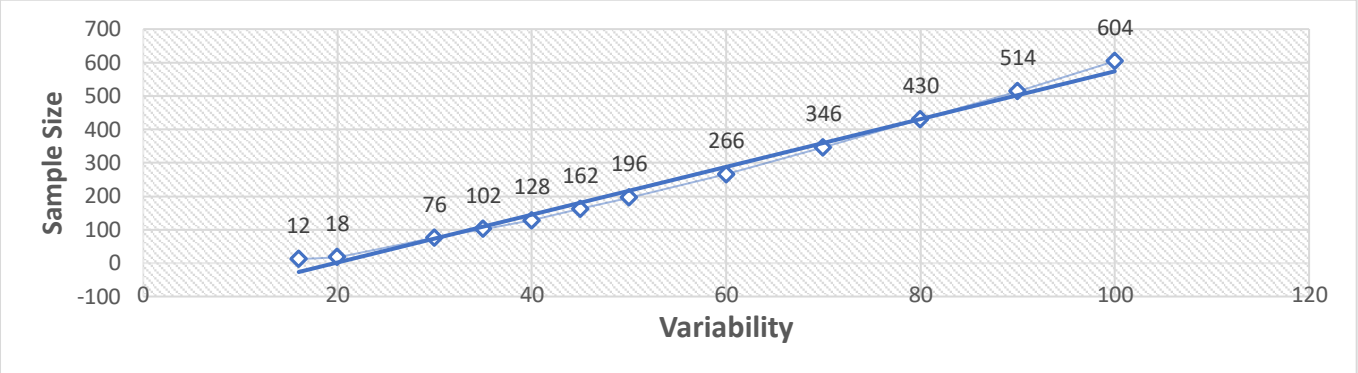
Table No. 6: Estimation of Sample Size based on the Equivalence Method:

Two Sample Mean, Two Arm Design- Parallel Group
Alpha = 0.05, Power = 0.80
Ho: 80 % > μ1/ μ2 > 100 %
H1: 0.80 % < μ1/ μ2 < 100 %

Lower Bound	Upper Bound	The ratio of Two Means	Variability (%)	Require Sample Size
80 %	120 %	95 %	20	40
80 %	120 %	95 %	25	60
80 %	120 %	95 %	30	82
80 %	120 %	95 %	40	140
80 %	120 %	95 %	50	208
80 %	120 %	95 %	60	284
80 %	120 %	95 %	70	368
80 %	120 %	95 %	75	412

Actual variability observed in the previous study was ~ 40 % between the High Pocket Expenditure and the Lower Pocket expenditure.

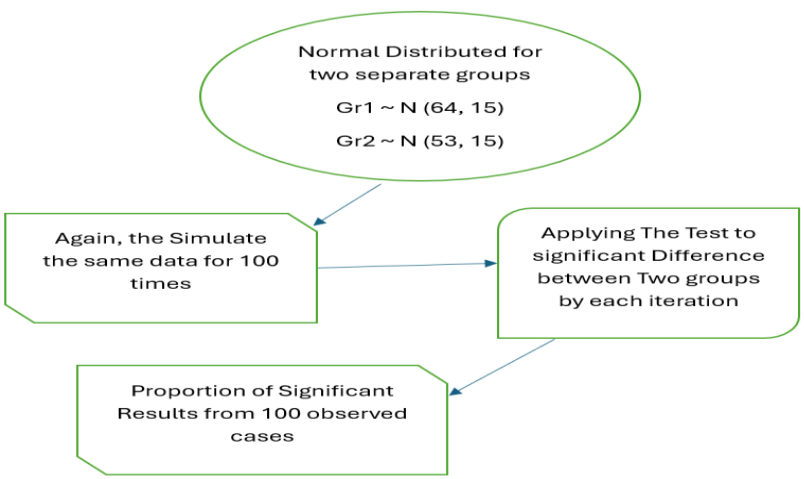
Figure #1: Two-Sample T-Test Procedure To Calculate The Sample Size For Ratio Of Mean [Variability Vs. Sample Size]



Simulation Method for Two-Sample T-Test [21, 22, 23]:

If researchers want to determine the sample size for a particular power level (like 80%), they can use a "deduce and verify" approach. However, a simulation-based technique might simplify the process of estimating the sample size. The study design of the clinical trial will include a description of the simulation strategy. The power of the sample size may be evaluated by simulating the data, doing statistical analysis on the simulated data, and calculating the ratio of significant results based on the statistical test. In this case, we used SAS software version 9.4 for the simulation techniques, which involve several steps —the first is simulating the data based on endpoint characteristics and design parameters. The second is running a statistical test on simulated data to see which cases are significant. In the last step, we can use the substantial case ratio as a power. Simulated the data with the Normal Distribution of Group #1 ~ N (64, 15) and a Normal Distribution of Group #2 ~ N (53, 15) with a sample size of 25. Again, the same sample size of set data was used with 100 simulations as described in Figure #2. According to simulations, assuming SD is equal in both groups, i.e., 15 units, and approximate mean of QoL scores are 64 and 53 in groups 1 & 2, respectively.

Figure #2: Flow Of Simulation Process Of Data



T-Test Results on Simulated Data

Simulated Data	t Value	Pr > |t|	Significant/ Non-Significant
1	4.9	<.0001	Significant
2	3.09	0.0033	Significant
3	4.2	0.0001	Significant
4	4.1	0.0002	Significant
5	2.21	0.0322	Significant
6	1.52	0.1359	Non-Significant
7	2.56	0.0137	Significant
8	1.3	0.2013	Non-Significant
9	4.13	0.0001	Significant
10	2.82	0.0069	Significant
11	5.76	<.0001	Significant
12	2.66	0.0105	Significant
13	2.63	0.0115	Significant
14	2	0.0509	Non-Significant
15	2.15	0.0367	Significant
16	2.67	0.0102	Significant
17	4.09	0.0002	Significant
18	3.86	0.0003	Significant
19	4.6	<.0001	Significant
20	2.28	0.0269	Significant
21	1.1	0.2762	Non-Significant
22	4	0.0002	Significant
23	3.93	0.0003	Significant
24	4.62	<.0001	Significant
25	3.9	0.0003	Significant
26	4.43	<.0001	Significant
27	3.95	0.0003	Significant
28	1.52	0.1344	Non-Significant
29	4.75	<.0001	Significant
30	3.01	0.0042	Significant
31	2.49	0.0164	Significant
32	2.16	0.0358	Significant
33	5.62	<.0001	Significant
34	3.61	0.0007	Significant
35	3.01	0.0042	Significant
36	4.25	<.0001	Significant
37	4.39	<.0001	Significant
38	3.64	0.0007	Significant
39	2.49	0.0164	Significant
40	4.2	0.0001	Significant
41	3.54	0.0009	Significant
42	2.63	0.0114	Significant
43	3.42	0.0013	Significant
44	4.81	<.0001	Significant
45	2.36	0.0224	Significant

46	3.46	0.0011	Significant
47	2.57	0.0134	Significant
48	2.53	0.0146	Significant
49	4.45	<.0001	Significant
50	3.61	0.0007	Significant
51	4.9	<.0001	Significant
52	1.06	0.2952	Non-Significant
53	1.48	0.1465	Non-Significant
54	3.86	0.0003	Significant
55	4.78	<.0001	Significant
56	1.63	0.1096	Non-Significant
57	3.87	0.0003	Significant
58	4.52	<.0001	Significant
59	2.97	0.0047	Significant
60	2.36	0.0224	Significant
61	3.81	0.0004	Significant
62	4.71	<.0001	Significant
63	2.26	0.0281	Significant
64	3.27	0.002	Significant
65	2.79	0.0076	Significant
66	3.53	0.0009	Significant
67	3.83	0.0004	Significant
68	3.84	0.0004	Significant
69	1.64	0.1085	Non-Significant
70	4.34	<.0001	Significant
71	5.13	<.0001	Significant
72	4.82	<.0001	Significant
73	1.26	0.215	Non-Significant
74	5.04	<.0001	Significant
75	3.28	0.0019	Significant
76	3.02	0.004	Significant
77	2.65	0.0108	Significant
78	5.41	<.0001	Significant
79	2.62	0.0119	Significant
80	3.34	0.0016	Significant
81	2.77	0.008	Significant
82	2.78	0.0077	Significant
83	3.57	0.0008	Significant
84	5.21	<.0001	Significant
85	3.93	0.0003	Significant
86	2.41	0.0199	Significant
87	3.41	0.0013	Significant
88	3.12	0.0031	Significant
89	3.25	0.0021	Significant
90	0.66	0.5095	Non-Significant
91	4.02	0.0002	Significant
92	1.43	0.1599	Non-Significant
93	2.93	0.0051	Significant
94	4.31	<.0001	Significant

95	2.82	0.0069	Significant
96	3.17	0.0027	Significant
97	2.78	0.0078	Significant
98	5	<.0001	Significant
99	2.31	0.025	Significant
100	2.9	0.0057	Significant

Table #7: Frequency of Significant Results From Simulated T-Test Results:

Proportion of "Significant"	
Proportion	0.88
95% Lower Confidence Limit	0.81
95% Upper Confidence Limit	0.94

From the above Table, there is an 88% probability of finding a 14-unit mean difference with a sample size of 25 in each group. This calculation is required to validate the statistical process in cases when exact prior knowledge is not accessible.

Discussion

This paper provides sample sizes for two scenarios, i.e., previous study results and simulation-based technique. The approach used to determine the sample size for clinical studies is based on the global health indicator endpoint that serves as the primary analytical focus. Sample Size analysis is a crucial tool for study planning, helping to balance Type I and Type II errors. It optimizes studies, improving detection of effects, saving money and time, and minimizing risks to subjects [22, 23, 24]. The standard statistical testing paradigm assumes type I errors are more critical than type II errors [23, 24, 25, 26]. Researchers are increasingly aware that past studies often had small sample sizes, leading to increased demand for larger samples in research. This may require more funding for participant payments or collaboration. Researchers can organize a collaborative study if a research question is essential but not feasible with current resources.

7.0 Recommendations:

Like other researchers, we saw a significant degree of heterogeneity in the mean differences in QoL score. However, the Minimal Clinically Important Difference should be guaranteed by the study's sample size justification; otherwise the study will not be meaningful. Therefore, a well-reasoned rationale for the justification of sample size is essential for the research. But sometimes prior information is not accurate in the current situation due to demographic or physical environment changes. In such cases, the simulation-based sample size estimation offers an alternative method for sample size determination, overcoming limitations of conventional formula-based approaches and ensuring sufficient power to detect clinically meaningful differences. It is flexible and can be easily implemented based on practical scenarios. However, a thorough exploration of data, based on objective evidence, is required in simulation. Then, the sample size can be approximated with real study data. This significant constraint arises during simulation, as researchers must address this issue during sample size computation.

REFERENCES

1. Phillips, A. and Haudiquet, V., 2003. ICH E9 guideline 'Statistical principles for clinical trials': a case study. *Statistics in medicine*, *22*(1), pp.1-11.
2. Das, S., Mitra, K. and Mandal, M., 2016. Sample size calculation: Basic principles. *Indian journal of anaesthesia*, *60*(9), pp.652-656.
3. KOTHAVALE, R.D., PATIL, Y.P., MAHAJAN, S.S. and GHUGE, Y.A., 2024. Sample Size Determination in Randomised Controlled Clinical Trials: A Narrative Review. *Journal of Clinical & Diagnostic Research*, *18*(8).
4. Michelson, H., Bolund, C., Nilsson, B. and Brandberg, Y., 2000. Health-related quality of life measured by the EORTC QLQ-C30: reference values from a large sample of the Swedish population. *Acta oncologica*, *39*(4), pp.477-484.
5. Coon, C.D., Schlichting, M. and Zhang, X., 2022. Interpreting within-patient changes on the EORTC QLQ-C30 and EORTC QLQ-LC13. *The Patient-Patient-Centered Outcomes Research*, *15*(6), pp.691-702.
6. Bergman, B., Aaronson, N.K., Ahmedzai, S., Kaasa, S. and Sullivan, M., 1994. The EORTC QLQ-LC13: a modular supplement to the EORTC core quality of life questionnaire (QLQ-C30) for use in lung cancer clinical trials. *European journal of cancer*, *30*(5), pp.635-642.
7. Makuch, R. and Simon, R., 1978. Sample size requirements for evaluating a conservative therapy. *Cancer Treat Rep*, *62*(7), pp.1037-1040.
8. Pocock, S.J., 2013. *Clinical trials: a practical approach*. John Wiley & Sons.
9. Edmiston Jr, C.E., Josephson, A., Pottinger, J., Ciacco-Tsivitis, M. and Palenik, C., 1993. The numbers game: sample-size determination. *American Journal of Infection Control*, *21*(3), pp.151-154.
10. Gail, M., 1973. The determination of sample sizes for trials involving several independent 2× 2 tables. *Journal of Chronic Diseases*, *26*(10), pp.669-673.
11. Freund, R.J. and Wilson, W.J., 2003. *Statistical methods*. Elsevier.
12. Horton, N.J. and Switzer, S.S., 2005. Statistical methods in the journal. *New England Journal of Medicine*, *353*(18), pp.1977-1979.
13. Makuch, R.W. and Simon, R.M., 1980. Sample size considerations for non-randomized comparative studies. *Journal of Chronic Diseases*, *33*(3), pp.175-181.
14. Gebert, P., Karsten, M.M., Hage, A.M., Dordevic, A.D. and Grittner, U., 2024. Statistical analysis plan for the PRO B study: open-label, superiority randomised controlled trial of alarm-based patient-reported outcome monitoring in patients with metastatic breast cancer. *Trials*, *25*(1), p.171.
15. Lachin, J.M., 1981. Introduction to sample size determination and power analysis for clinical trials. *Controlled clinical trials*, *2*(2), pp.93-113.
16. Simon, R., 1987. How Large Should a Phase II Trial of a New Drug Be? 1. *Cancer treatment reports*, *71*(11), p.1079.
17. Gross, A.J., Hunt, H.H., Cantor, A.B. and Clark, B.C., 1987. Sample size determination in clinical trials with an emphasis on exponentially distributed responses. *Biometrics*, pp.875-883.
18. Makuch, R., 1985. Approaches to sample size estimation in the design of clinical trials--a review. *Statistics in Medicine*, *4*(2), pp.247-247.
19. Castelloe, J.M., 2000, April. Sample size computations and power analysis with the SAS system. In *Proceedings of the Twenty-Fifth Annual SAS User's Group International Conference* (pp. 265-25).
20. Cody, R., 2011. *SAS statistics by example*. SAS Institute.
21. Edition, T., 2014. Base SAS® 9.4 Procedures Guide.
22. Lenth, R.V., 2001. Some practical guidelines for effective sample size determination. *The american statistician*, *55*(3), pp.187-193.
23. Muller, K.E. and Benignus, V.A., 1992. Increasing scientific power with statistical power. *Neurotoxicology and teratology*, *14*(3), pp.211-219.
24. O'Brien, R.G. and Castelloe, J., 2007. Sample size analysis for traditional hypothesis testing: concepts and issues. *Pharmaceutical statistics using SAS: A practical guide*, pp.237-71.
25. Peters, J.R., 2014. Clinical endpoint bioequivalence study. In *FDA Bioequivalence Standards* (pp. 243-274). New York, NY: Springer, New York.
26. Hauschke, D., 2002. A note on sample size calculation in bioequivalence trials. *Journal of pharmacokinetics and pharmacodynamics*, *29*, pp.89-94.