KEY POINTS IN SAMPLE SIZE WORKSHOP

By: Nur Aazifah[1], Adam Bujang[2], Nurul Iman Hafizah[1], Boon-How Chew[1]

[1]Cllinical Research Unit, HPUPM
[2]Clinical Research Centre, Hospital Umum Sarawak

Justification of sample size calculation is a vital part of any clinical research. However, estimating the number of participants required to give a valid result is not always easy. Studies that have a sample size that is too small will be underpowered and may lead to inconclusive results, while too large of sample size will lead to waste of resources and expose more participants than necessary to any related risk related to the study. Important components that are required in the calculation of sample size include study design, an estimated important effect size, type 1 error, type 2 error, desired power, also sometime number of variables and precision are relevant considerations. A brief discussion of the important components of sample size calculation had been discussed in a previous article on Determination of Sample Size (CLICK HERE).

This article will briefly describe important steps in sample size calculation for clinical trials followed with observational studies. In addition, this article will introduce the methods to calculate sample size required for studies using common statistical analysis in multivariable modelling. There are various ways to estimate sample size required for a proposed study.

Sample size calculation can be done manually using specific formula or sample size software can be used to ease the calculation. Common free software to calculate sample size is available at Software & Calculator | HOSPITAL PENGAJAR UNIVERSITI PUTRA MALAYSIA (upm.edu.my). Nowadays, scholars have tabulated sample size table from various statistical test and these are also available in the literatures. Some scholars presented their sample size estimation using nomograms."

Study Design 1: Randomised Controlled Trial

Randomised controlled trials (RCT) are prospective studies that commonly used to measure the effectiveness of a new intervention or treatment. Many clinical trials that do not carefully consider the sample size requirement turn out to lack the statistical power or the ability to detect intervention effects of magnitude that has clinical importance (45,46). The numerous designs of RCT such as parallel RCT, cluster RCT, and factorial will require slightly different sample size estimation approaches. This article will demonstrate an example for parallel RCT which is the most common RCT. The method to calculate sample size for other design of RCT will be discussed in future articles.

Generally, there are **two types of formula to calculate sample size in RCT** which are **two proportion** which be used in **dichotomous data** (the outcome) and **two means** which being used for **continuous variable** (the outcome) with the assumption that the sample are recruited and assigned randomly to the groups.

Using the recently published study on JAMA, the ITECH trial with the aim to determine the efficacy of ivermectin in preventing progression to severe disease among high-risk patients with COVID-19.

Let's go through the statement in study on the section of sample size.

The sample size was calculated based on a superiority trial design and primary outcome measure. The expected rate of primary outcome was **17.5%** in the control group, according to a previous local data of high-risk patients who presented with mild to moderate disease. **A 50% reduction of primary outcome**, or a 9% rate difference between intervention and control groups, was considered clinically important. This trial required 462 patients to be adequately powered. This sample size provided a **level of significance at 5% with 80% power for 2-sided tests**. Considering potential dropouts, a total of 500 patients (250 patients for each group) were recruited.

1) Using two proportion formula (Pocok's formula)

$n = [(p_1 (1-p_1) + p_2 (1-p_2)] x (Z_\alpha + Z_\beta)^2 / (p_1-p_2)^2$

$n = \frac{[0.175(1-0.175)+0.087(1-0.087)] x (0.84+1.96)^2}{(0.175-0.087)^2}$

= 228 per arm, so there are two arms in the trial

Total sample size required: 457 with 10% drop up, round up to nearest number total 500 participants needed.

where :

n = required sample size
α = level of statistical significance
1-β = power of study
zα = value of the standard normal distribution cutting off probability α in one tail for a one –sided alternative or α/2 in each tail for a twosided alternative
zβ = value of the standard normal distribution cutting off probability β

2) Using G-power software using exact test

Exact – Proportions: Inequality, two independent groups (Fisher's exact test)

Options:	Exact distribution		
Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Proportion p1	=	0.1750000
	Proportion p2	=	0.08699999999999999
	α err prob	=	0.05
	Power (1-β err prob)	=	0.8
	Allocation ratio N2/N1	=	1
Output:	Sample size group 1	=	247
	Sample size group 2	=	247
	Total sample size	=	**494**
	Actual power	=	0.8011141
	Actual α	=	0.0352500

	Z_α	
α	One-sided test	Two-sided test
0.10	1.282	1.645
0.05	1.645	1.960
0.025	1.960	2.240
0.01	2.326	2.576

$1-\beta$	Z_β
0.50	0.00
0.60	0.25
0.70	0.53
0.80	0.84
0.85	1.036
0.90	1.282
0.95	1.645
0.975	1.960
0.99	2.326

The concept of sample proportion as shown above is relevant, however, modification is needed to calculate sample size for **continuous variable**. Below is the formula for continuous outcome variable:

$N = \frac{2\sigma^2 (Z\alpha+Z\beta)^2}{\Delta^2}$

Where:
σ = standard deviation of either group
Δ = expected detectable difference between two groups
zα = value of the standard normal distribution cutting off probability α in one tail for a one –sided alternative or α/2 in each tail for a two-sided alternative

Below is an example of sample size calculation for continuous outcome variable which is blood pressure.

A new antihypertensive drug is to be tested against current treatment practice in people with systolic blood pressure > 160 mmHg and/or diastolic blood pressure > 95mmHg. It is felt that if the new drug can achieve blood pressure levels that are on the average **10 mmHg** than those achieve using current treatment then it would be accepted by the medical community. The investigators would like at **least 90% power** and have chosen **α = 0.01 (two-sided)** as the current therapy is quite acceptable and they want to be sure that the new therapy is superior before switching over. Blood pressure measurement has a standard deviation of **20 mmHg.**

α= 0.01 Δ=10 Z α=2.58 β= 0.1 σ=20 Z β=1.28

Substitute the value into the formula

$N = \frac{2\sigma^2 (Z\alpha+Z\beta)^2}{\Delta^2}$

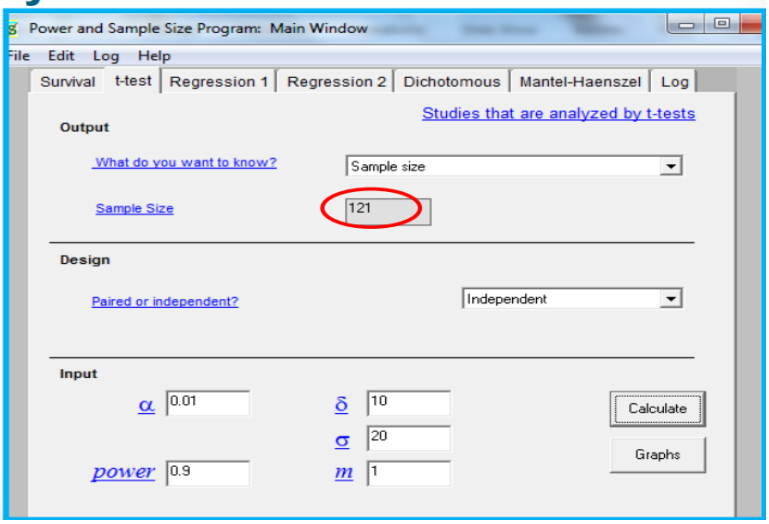
n = 119.2

Required sample size is **120** per group (240 hypertensive in all)
It is recommended to consider 10 -20% drop out rates in the sample size calculation.

The number of sample size calculated using the software also yielded almost similar value which is **121** subjects per group.

On the other hand, the figure below showed the calculation of sample size using the software Power and Sample Size.

By software



Study Design 2: Observational Study

Cohort, cross sectional, and case-control studies are examples of data collection designs in observational studies. Often, these studies are the only practicable method of studying various problems related to a disease of interest, for example, studies of aetiology are one of the instances where a randomised controlled trial might be unethical, or if the condition to be studied is rare.

Researchers can utilise similar formula to calculate sample size. However, little modification is needed for calculation sample descriptive studies which mainly aimed to determine the prevalence of diseases size. One proportion sample size formula can be used to calculate sample size in descriptive studies. The main difference between one proportion and two proportion formulae is the calculation in one proportion formula do not involve hypothesis testing thus power is not included in the formula.

One proportion sample size formula:

$$n = \left(\frac{z}{\Delta}\right)^2 p(1-p)$$

Where:

p : expected proportion of individuals in the sample with the characteristic of interest at the determined 100(1-α)% confidence interval. It can be obtained from literature or a pilot study or preliminary work

Δ = precision (generally at 0.05, however it can be adjustable to achieve affordable, feasible and statistically meaningful sample size

Below is an example of sample size calculation using one proportion formula:

A local health department wishes to estimate the prevalence of dental carries among children under 12 years of age in its locality. How many children should be included in the sample so that it may be estimated to within **5 percentage points of the true value with 95% confidence**? It has been estimated that the prevalence of dental carries among children was **20%** from previous literature

Solution :
Anticipated population proportion (p) = 20% (0.2)
Level of significance = 5% (0.05)
Absolute precision (Δ) = $\pm$ 5%
n = 246

The sample of 246 children required at the analysis stage.

3) Calculation Based on Statistical Analysis:

Multivariate analysis deals with simultaneously predicting **multiple** outcomes while multivariable analysis is a tool for determining the relative contributions of different factors to a **single** event.

Observational study that is causal in nature will has many confounding factors that can be controlled using multivariable analyses. Generally, the number of sample size required for observational studies with planned multivariable analysis is higher compared to univariate and bivariate analysis. The number of sample size is heavily depended on the number of independent variables in the final model.

Different types of statistical test require different method of sample size calculation.

Table 1 shows the published articles related to sample size determination for various statistical tests.

Statistical test	Published articles
a/ To estimate parameters for population	Krejcie and Morgan (1), Lachin (2), Campbell et al. (3), Bartlett et al. (4), Israel (6), Naing et al. (7)
b/To infer the results for larger population	
Correlation	Cohen (8), Algina and Olejnik (9), Bujang and Nurakmal (10)
Intra-class correlation	Fleiss and Cohen (11), Bonett (12), Zou (13), Bujang and Baharum (14)
Kappa agreement test	Cicchetti (15), Flack et al. (16), Cantor (17), Sim and Wright (18), Bujang and Baharum (19)
Independent sample t-test and paired t-test	Lachin (2), Cohen (8), Dupont and Plummer (20).
One-way ANOVA	Cohen (8), Jan and Shieh (21)
Pearson's chi-square	Lachin (2), Cohen (8), Dupont and Plummer (20)
Cronbach's alpha	Bonett (22), Bonett (23), Bonett and Wright (24), Bujang et al.(25)
Sensitivity and specificity	Buderer (26), Malhotra and Indrayan (27), Bujang and Adnan (28)
Linear regression or Multiple linear regression	Cohen (8), Dupont and Plummer (20), Hsieh et al. (29), Knofczynski and Mundfrom (30), Tabachnick and Fidell (31), Bujang et al. (32).
Analysis of covariance	Borm et al. (33), Bujang et al. (34)
Logistic regression	Peduzzi et al. (35), Hsieh et al. (29), Bujang et al. (34)
Survival analysis	Lachin (2), Lachin and Foulkes (36), Dupont and Plummer (20).
Cox regression	Peduzzi et al. (37), Hsieh and Lavori (38), Schmoor et al. (39).
Exploratory factor analysis	Barrett and Kline (40), Osborne and Costello (41), Bujang et al. (42),Bujang et al. (43).

Table adapted from Bujang MA. A step-by-step process on sample size determination for medical research Malays J Med Sci. 2021;28(2):15–27. https://doi.org/10.21315/mjms2021.28.2.2.

Next, this article will discuss on the rule of thumb for common statistical test used in medical and clinical research which include logistic regression, cox regression, multiple linear regression and analysis of covariance (ANCOVA).

i) Logistic regression and cox regression:

The similarities between logistic regression and cox regression are both have binary outcome. Therefore, similar formula can be applied to calculate sample size. Previous study by Peduzzini et al (1996) suggested to used EPV 10 (event per variable = 10) where the rule of thumb depends on a few parameters which are:
1/ Prevalence of the outcome of interest
2/ Number of participants to be recruited
3/ Number of independent risk factor on final model

However, the rule received some critics and recommended to used EPV20 instead of EPV50. In a latest publication by Bujang et al (2018), the author recommend a simplified version of formula which is : **n = 100 + 5i** where i refers to number of independent variables in the final regression model.

ii) Multiple linear regression (MLR) and analysis of covariance (ANCOVA)

MLR and ANCOVA share a common assumption however usually applied in different scenario. The proposed formula to be used in multiple linear regression (MLR) and general linear model (ANCOVA) is **N> 50 + 8M** as proposed by Tabachanick et al (2013).

Where :
N = sample size required
M= no of predictors or risk factor

Although sample size estimation based on a rule of thumb may considered as a weak method compare to the proper sample size calculation, but scholars have proposed rule of thumbs to ease researchers. The idea is researchers to be able to come out with sufficient sample size that will likely prevent the study from underpowered and at the same time prevent them from wasting resources. In addition, it is not practical to calculate sample when the minimally important effect sizes are unknown and unpredictable.

In a nutshell, there is no one-size-fits-all formula for sample size calculation that will be able to fit all study designs and statistical analyses. Sample size must be calculation properly to ensure the study have enough power to justified the aim of the study.

References :

1. Krejcie RV, Morgan DW. Determining sample size for research activities. Educ Psychol Meas. 1970;30:607–610. https://doi.org/10.1177/001316447003000308
2. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. Controlled Clin Trials. 1981;2(2):93–113. https://doi.org/10.1016/0197-2456(81)90001-5
3. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. BMJ. 1995;311:1145–1148. https://doi. org/10.1136/bmj.311.7013.1145
4. Bartlett JE, Kotrlik JW, Higgins C. Organizational research: determining appropriate sample size for survey research. Inf Technol Learn Perform J. 2001;19:43–50
5. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85:257–268. https://doi.org/10.1093/ptj/85.3.257
6. Israel GD. Determining sample size (Tech. Rep. No. PEOD-6). Florida: University of Florida, Institute of Food and Agricultural Sciences; 2003
7. Naing L, Winn T, Rusli BN. Practical issues in calculating the sample size for prevalence studies. Archives of Orofacial Sciences. 2006;1:9–14.
8. Cohen J. A power primer. Psychol Bull. 1992;112:155–159. https://doi.org/10.1037/0033- 2909.112.1.155
9. Algina J, Olejnik S. Sample size tables for correlation analysis with applications in partial correlation and multiple regression analysis. Multivar Behav Res. 2003;38:309–323. https:// doi.org/10.1207/S15327906MBR3803_02
10. Bujang MA, Nurakmal B. Sample size guideline for correlation analysis. World Journal of Social Science Research. 2016;3(1):37–46. https://doi. org/10.22158/wjssr.v3n1p37

11. Fleiss J, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Edu Psychol Meas. 1973;33:613–619. https://doi. org/10.1177/001316447303300309
12. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med. 2002;21:1331–1335. https:// doi.org/10.1002/sim.1108 25.
13. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med. 2012;31(29):3972– 3981. https://doi.org/10.1002/sim.5466 .
14. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch Orofac Sci. 2017;12:1– 11.
15. Cicchetti DV. Testing the normal approximation and minimal sample size requirements of weighted kappa when the number of categories is large. Appl Psychol Meas. 1981;5(1):101–104. https://doi.org/10.1177/014662168100500114
16. Flack V, Afifi A, Lachenbruch P. Sample size determinations for the two rater kappa statistic. Psychometrika. 1988;53:321–325. https://doi. org/10.1007/BF02294215
17. Cantor AB. Sample-size calculations for Cohen's kappa. Psychol Methods. 1996;1(2):150–153. https://doi.org/10.1037/1082-989X.1.2.150
18. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85:257–268. https://doi.org/10.1093/ptj/85.3.257
19. Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Kappa agreement test. Epidemiology, Biostatistics, and Public Health. 2017;14(2);e12267-1. https://doi. org/10.2427/12267
20. Dupont WD, Plummer WD. Power and sample size calculations for studies involving linear regression. Control Clin Trials. 1998;19:589–601. https://doi.org/10.1016/S0197-2456(98)00037-3
21. Jan SL, Shieh G. Sample size determinations for Welch's test in one-way heteroscedastic ANOVA. Br J Math Psychol. 2014;67:72–93. https://doi. org/10.1111/bmsp.12006
22. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. J Educ Behav Stat. 2002;27:335–340. https://doi. org/10.3102/10769986027004335
23. Bonett DG. Sample size requirements for comparing two alpha coefficients. Appl Psychol Meas. 2003;27(1):72–74. https://doi. org/10.1177/0146621602239477 35.
24. Bonett DG, Wright TA. Cronbach's alpha reliability: interval estimation, hypothesis testing, and sample size planning. J Organ Behav. 2015;36(1):3–15. https://doi.org/10.1002/ job.1960 36.
25. Bujang MA, Omar ED, Baharum NA. A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. Malays J Med Sci, 2018;25(6):85–99. https://doi.org/10.21315/ mjms2018.25.6.9
26. Buderer NM. Statistical methodology: incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med. 1996;3:895– 900. https://doi.org/10.1111/j.1553-2712.1996. tb03538.x 38.
27. Malhotra RK, Indrayan A. Simple nomogram for estimating sample size for sensitivity and specificity of medical tests. Indian J Ophthalmol. 2010;58:519–522. https://doi.org/10.4103/0301- 4738.71699
28. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. J Clin Diagn Res. 2016;10(10):YE01–YE06. https://doi.org/10 .7860/JCDR/2016/18129.8744
29. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. Stat Med. 1998;17(14):1623– 1634. https://doi.org/10.1002/(SICI)1097- 0258(19980730)17:14%3C1623::AID-SIM871%3E 3.0.CO;2-S
30. Knofczynski GT, Mundfrom D. Sample sizes when using multiple linear regression for prediction. Educ Psychol Meas. 2008;68(3):431–442. https://doi.org/10.1177/0013164407310131
31. Tabachnick BG, Fidell LS. Using multivariate statistics. 6th ed. Boston: Pearson Education; 2013.
32. Bujang MA, Sa'at N, Tg Abu Bakar Sidik TMI. Determination of minimum sample size requirement for multiple linear regression and analysis of covariance based on experimental and non-experimental studies. Epidemiology Biostatistics and Public Health. 2017;14(3):e12117-1. https://doi. org/10.2427/1211
33. Borm GF, Fransen J, Lemmens WA. A simple sample size formula for analysis of covariance in randomized clinical trials. J Clin Epidemiol. 2007;60:1234–1238. https://doi.org/10.1016/j. jclinepi.2007.02.006
34. Bujang MA, Sa'at N, Tg Abu Bakar Sidik TMI, Lim CH. Sample size guidelines for logistic regression from observational studies with large population: emphasis on the accuracy between statistics and parameters based on real life clinical data. Malays J Med Sci. 2018;25(4):122–130. https://doi. org/10.21315/mjms2018.25.4.12
35. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996:49;1373–1379. https:// doi.org/10.1016/S0895-4356(96)00236-3
36. Lachin JM, Foulkes MA. Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. Biometrics. 1986;42:507–519. https://doi. org/10.2307/2531201
37. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol. 1995;48:1503–1510. https://doi. org/10.1016/0895-4356(95)00048-8
38. Hsieh FY, Lavori PW. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. Control Clin Trials. 2000;21:552–560. https://doi.org/10.1016/ S0197-2456(00)00104-5
39. Schmoor C, Sauerbrei W, Schumacher M. Sample size considerations for the evaluation of prognostic factors in survival analysis. Stat Med. 2000;19:441–452. https://doi.org/10.1002/ (SICI)1097-0258(20000229)19:4%3C441::AIDSIM349%3E3.0.CO;2-N
40. Barrett PT, Kline P. The observation to variable ratio in factor analysis. Personality Study in Group Behavior. 1981;1:23–33.
41. Osborne J, Costello A. Sample size and subject to item ratio in principal components analysis. Practical Assessment, Research & Evaluation. 2004;9(11):Article 11. https://doi.org/10.7275/ ktzq-jq66
42. Bujang MA, Ab Ghani P, Soelar SA, Zulkifli NA. Sample size guideline for exploratory factor analysis when using small sample: taking into considerations of different measurement scales. Statistics in Science, Business, and Engineering (ICSSBE) 2012, Langkawi, Malaysia, 10–12 September 2012. https://doi.org/10.1109/ ICSSBE.2012.6396605
43. Bujang MA, Ghani PA, Soelar SA, Zulkifli NA, Omar ED. Invalid skewed responses contributes to invalid factor solution in exploratory factor analysis: a validation approach using real-life data. J Behav Health. 2019;8(4):152–160. https://doi.org/10.5455/jbh.20190628084939
44. Naing et al Determination of Sample size: Concept and Application, Determination of Sample Size in Interventional Study, Determination of Sample size in Analytical Study, Determination of Sample size in descriptive study Sample Size Workshop Lecture Notes Feb 2022
45. Freiman JA, Chalmers TC, Smith H, Jr., Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. N Engl J Med 1978;299:690–694.
46. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122–124