Column: Tips on Research and Publication

SAMPLE SIZE ESTIMATION IN RESEARCH: NECESSITY OR COMPROMISE?

Samir Kumar Praharaj*[1], Shahul Ameen[1]

[1]Department of Psychiatry, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal, Karnataka, India
[2]Consultant Psychiatrist, St. Thomas Hospital, Changanacherry, Kerala, India
*Corresponding address: Professor & Head, Department of Psychiatry, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal, Karnataka, India, PIN- 576104. Email address: samirpsyche@yahoo.co.in

Abstract

An adequately powered sample is essential for accurate parameter estimation and meaningful significance testing. It is important to balance the sample size with practical considerations such as cost and feasibility. Sample size calculation is guided by key factors such as effect size, variability, power, and significance levels. While complex formulas and software aid precision, practical rules of thumb and strategies can also be used effectively. Transparent documentation of the rationale and methods used for sample size calculation is vital for ensuring reproducibility.

Most studies aim to estimate a 'parameter' in a population based on a 'point estimate' from the sample. To do this effectively, two things are essential:

a) The sample should be *representative* and
b) The sample should be *adequate*.

This paper focuses on the latter. Sample size estimation is required to correctly identify the adequate sample for a study. Determining the sample size is essential to ensure:

a) Hypothesis testing: To minimize the chance of missing an effect if it truly exists (i.e., adequate power) and
b) Parameter estimation: To determine the range of the estimate (i.e., confidence interval).

However, many studies, including those in psychiatry, are often underpowered. For instance, Califf et al. (2012) reported a median of 61 participants per study in mental health clinical trials.[1] Nearly half of all studies did not adequately plan for sample size.[2]

In any study, a larger sample provides a more precise estimate. However, it can be expensive, impractical, and sometimes ethically questionable if the risks outweigh the benefits. On the other hand, a small sample may not yield reliable results and might lack the power to detect true effects. *Post hoc* power calculations are also generally not useful as they do not provide any additional information than what is available from the P value.[3] Therefore, it is recommended to conduct *a priori* sample size estimation for all quantitative studies. Qualitative study designs, however, do not formally require sample size estimation. Instead, they rely on the principle of data saturation, where sampling continues

Access this article online:
https://kjponline.com/index.php/kjp/article/view/463
DOI: 10.30834/KJP.37.1.2024.463
Received on: 11/07/2024. Accepted on: 29/08/2024.
Web publication: 30/08/2024

QR code:

Please cite this article as: Praharaj SK, Ameen S. Sample size estimation in research: Necessity or compromise? Kerala Journal of Psychiatry 2024;37(1):66-71.

until no new information emerges, at which point data collection can stop.

Statistical Concepts Related to Sample Size

The right sample size typically hinges on specific statistical hypotheses and key study design factors. These include the smallest detectable difference that is meaningful (effect size), estimated variability in measurements, desired statistical power, and significance level. An understanding of these variables and their interplay in estimating sample size is crucial for study planning.

1. *Outcome measure*: The sample size calculation depends on the level of measurement, such as binary or continuous. Continuous outcomes typically require a smaller sample size compared to categorical outcomes.
2. *Applied statistical test***:** The planned statistical test, such as a t-test or chi-squared test, influences the calculation of sample size.
3. *α level*: This represents the probability of type I error, typically set at less than 5%. Other cut-offs (e.g., 1%) can also be used. Using a lower α level increases the required sample size.
4. *β level*: This is the probability of type II error, usually set at 20%. A lower β level necessitates a higher sample size.
5. *Statistical power*: This complements type II error (1 – β). It is typically set at 80%, but higher values can be chosen. A higher power requires a larger sample size.
6. *Number of conditions*: A study with more groups or conditions requires a larger sample size.
7. *Expected treatment effect*: The magnitude of the effect, e.g., incidence or effect size, impacts the sample size. A larger effect size reduces the necessary sample size.
8. *Number of repeated measures*: While repeated measures designs generally need a smaller sample size, the required sample size increases as the number of repeated measures increases.
9. *Correlation among repeated measures*: A higher correlation among variables in a repeated measures study design can increase the necessary sample size. Since the variability between repeated measurements is reduced, a larger sample size may be required to detect a meaningful effect.
10. *Expected dropout*: Estimated dropout rates, typically around 10-20%, should be considered in sample size calculations. Higher dropout rates, such as in studies involving participants with substance use disorder, may require adjustments to sample size estimates.
11. *Population size*: If the study is conducted in a finite population, the required sample size is less than the larger population.

Besides the above-mentioned factors, other contextual considerations for sample size estimation include:

1. *Funding*: Non-funded studies may necessitate a smaller sample size due to financial constraints.
2. *Feasibility*: Pragmatic reasons, such as limited study duration, may necessitate a smaller sample size.
3. *Ease of access to sample*: Challenges in accessing a sample in a specific setting may limit the ability to recruit more participants.
4. *Limited scalability*: Preliminary studies with constrained resources may justify smaller sample sizes.

However, using smaller samples than necessary compromises the quality of the study.

Basic Steps for *a priori* Sample Size Estimation

The following steps are common for most of the studies:

1. *Define the study population*: Determine the population to which the study findings will apply. Identify the accessible population from which the sample will be drawn.

Table 1. Sample size thumb rules

Sl No	Thumb rules or formulae	Description
1	Roscoe's rule of thumb	a) Sample size >30 and <500 is sufficient for most research b) Comparison groups should have 30 participants each c) Sample size is ten times the number of variables d) A small sample of 10 or 20 will suffice for new experiments
2	Lehr's formula	Sample size = $16/(\text{standardized difference})^2$ Comparison of two equal-sized groups, 80% power, a two-tailed significance level of 0.05
3	Krejcie and Morgan's table	A sample size of 384 is sufficient for a population of 1,000,000 or more (the sample should be representative)
4	Green's procedures for regression analysis	$N \geq 50+8m$ (where m refers to the number of predictors in the model)
5	Rule of 100 for EFA (Other rules are 150, 200, 250, 300, and 500)	Minimum 100 samples even though the number of variables is less than 20 (Gorsuch 1983) The number of subjects should be the larger of 5 times the number of variables, or 100 (Hatcher 1994)
6	Barclay's 10-times rule for PLS-SEM	The minimum sample size should be: a) 10 times the number of indicators to measure one construct b) 10 times the number of structural paths directed at a particular latent construct
6	Kline's sample size guidelines for SEM	Less than 100 is inadequate, 100 is small, 100 to 200 is medium, and over 200 is large
7	Kreft's '30/30' rule for multilevel models	Minimum of 30 groups with 30 individuals per group (alternative Hox's 50/20 rule, i.e. 50 groups with 20 per group)
8	Nunnally's guideline to cross-validate results of regression analysis	To select the best variables from 10 possible ones, 400 to 500 samples is required

EFA: Exploratory factor analysis; PLS-SEM: Partial least square – structural equation modeling

2. *Choose the study design*: Select a design that aligns with the research question and objectives.
3. *Specify hypothesis, significance level, and power*: Clearly outline the primary objective that will drive the sample size calculation. Set the significance level at 0.05 and aim for a statistical power of 80%.
4. *Gather relevant information about the parameters*: Define the outcome measure (e.g., means or proportions) clearly. Use only the primary outcome for sample size estimation. Obtain the effect size from existing literature or make a preliminary estimate based on clinical expertise or a pilot study if prior data is unavailable.
5. *Calculate sample size using the parameters and choose an appropriate one for the study*: Use the specified parameters to calculate the sample size, considering reasonable variations. For example, the sample sizes should be calculated for 80% and 90% power, and the higher power level for the study should be chosen if feasible.
6. *Adjust the sample size based on estimated dropout rates*: Account for potential participant dropout by increasing the sample size, typically by 10 to 20%, to ensure adequate statistical power and account for attrition during the study.

How to Determine Effect Size?

The magnitude of the effect is required to calculate the sample size. The chosen effect size must be clinically significant and meaningful; hence, it involves some judgment

Table 2: Sample size estimation tools

	Name	Link
Online calculators	Riskcalc	https://riskcalc.org/samplesize/
	Epitools	https://epitools.ausvet.com.au/samplesize
	Openepi	https://www.openepi.com/SampleSize/SSPropor.htm
	ClinCalc	https://clincalc.com/stats/samplesize.aspx
	Sample-Size	https://sample-size.net/
	EasyMedStat	https://www.easymedstat.com/sample-size-calculator
	Statulator	https://statulator.com/SampleSize/ss1P.html
Excel files	Scalex SP	https://sites.google.com/view/sr-ln/ssc
	Sample Size Calculator v2.0	https://wnarifin.github.io/ssc/Sample%20Size%20Calculator%20v2.0.xls
	Sample Size Estimation and Power Calculations	https://www.ucl.ac.uk/child-health/short-courses-events/about-statistical-courses/sample-size-estimation-and-power-calculations
Software	G*Power	https://download.cnet.com/g-power/3001-2054_4-10647044.html
	PASS software	https://www.ncss.com/software/pass/
	nQuery	https://www.statsols.com/
	GLIMMPSE	https://glimmpse.samplesizeshop.org/
	XLSTAT	https://www.xlstat.com/en/download

from the researcher. The following methods may be considered:

1. *Prior experience* or *judgment*: Use insights gained from previous studies or clinical experience to estimate the magnitude of the effect.
2. *Meta-analyses* and *well-designed studies*: Refer to meta-analyses for pooled estimates from multiple studies or consider findings from well-designed individual studies if meta-analyses are unavailable.
3. *Pilot studies*: If no prior research is available, conduct pilot studies. For the pilot studies, no sample size calculation is required.
4. *Standardized effect size measures*: Consider using standardized effect size measures, categorized as small, medium, or large effects. While somewhat arbitrary, these measures provide a framework for estimating effect sizes in studies.

How to Calculate Sample Size?

Sample size calculations typically involve using specific formulas tailored to the study design. Readers can consult comprehensive sources like Machin et al. (2018) for detailed methods of calculating sample size across various study designs.[4] However, some researchers argue that complex formulas may not always be necessary, and simple rules of thumb can often suffice. Roscoe (1975) proposed several such rules:

a) For most research, a sample size between 30 and 500 is generally sufficient
b) In comparison studies, each group should have at least 30 participants based on the central limit theorem
c) For multivariate analysis, the sample size should be at least ten times the number of variables, and
d) Simple experimental research can be done with smaller sample sizes, such as 10 to 20 participants.

These rules provide practical guidelines for estimating sample sizes without complex calculations.[5] Altman's nomogram, detailed in O'Hara (2008), offers a useful tool for estimating sample size based on standardized differences and desired statistical power. For a list of such thumb rules and simple formulas, see Table 1.[6]

There are numerous tools available to aid in

calculating sample size, including online calculators and Excel sheets (refer to Table 2 for examples). Moreover, various free and paid software options have been developed specifically for this purpose. G*Power is a widely used, user-friendly, free software that helps calculate sample size across a range of study designs. However, it is essential to recognize that for complex designs, such as multivariate studies, the sample size estimates provided by G*Power may not be very accurate. Researchers should consider consulting specialized statistical software or seeking guidance from statisticians when dealing with intricate study designs to ensure accurate sample size calculations.

Strategies for Reducing Sample Size

When conducting studies where large samples are not practical or feasible, researchers may need to adopt methods to adjust sample size, albeit with some compromise in study quality. Here are some approaches that can be considered:

1. *Reduce statistical power*: Lowering the desired statistical power, for example, from 90% to 80%, decreases the required sample size. However, this increases the risk of not detecting true effects if they exist.
2. *Use continuous outcomes*: Preferentially selecting continuous variables as primary outcome measures reduces the required sample size compared to dichotomous outcomes. Additionally, continuous measures are more likely to yield significant findings than categorical outcomes.
3. *Enrich the population*: Smaller sample sizes may be sufficient if the study population is homogeneous with reduced variability. However, narrowing eligibility criteria limits the generalizability of findings beyond the specific population studied.
4. *Use repeated measurements*: Study designs with repeated measurements can reduce the required sample size, but the choice of design ultimately depends on the research question.
5. *Reduce the dropout rate*: Minimize participant dropout through strategies like regular follow-ups, frequent reminders, incentives, and ensuring participant engagement.
6. *Use unequal group sizes*: While equal group sizes are ideal for ensuring statistical power, sometimes unequal group sizes (e.g., 2:1 or 3:1 ratio) can be considered. This approach may be justified if one group is more difficult to recruit or retain.
7. *Increase effect size estimate*: The effect size estimate is usually based on experience or prior research. However, considering medium to large effect size may be justified for preliminary research. The preliminary study findings can inform the appropriate sample size for later studies.
8. *Use surrogate outcomes*: Using surrogate outcomes that correlate with the primary outcome can potentially reduce the required sample size. Surrogate outcomes are often easier or quicker to measure than the primary outcome and may be used if a larger sample size is not feasible. For example, sleep abnormalities in depression or sensory gating abnormalities for schizophrenia can be considered as surrogate outcomes.[7] Similarly, quality of life measures can be a surrogate marker for depression symptoms.
9. *Increase the event rate*: To increase the event rate and thus potentially decrease the necessary sample size, extend the follow-up duration, use surrogate outcomes, or consider composite outcomes that combine multiple outcomes. Some examples of composite measures are body-mass index, intelligence quotient, and WHO quality of life score.

How Much to Report?

The method for determining sample size in a study should be described in sufficient detail to allow its use in other protocols. Key elements

that should be clearly documented include the power, significance level, mean or rate for the control group, minimal detectable difference, variance, and dropout rate. Moreover, any other factors that form the basis of sample size calculation should also be included. It is essential to cite relevant prior studies or meta-analyses that provided effect size measures used in the determination. Specifically, the documentation should cover:

a) The planned sample size based on the calculation,
b) The actual sample size achieved, noting any discrepancies from the planned size,
c) Details of how the sample size was determined, including methodology for power analysis and effect size estimation, and
d) Any planned interim analysis and stopping rules applied during the study.

Last Remarks

Sample size estimation is crucial for all studies as it enhances the internal validity of the research. Large and small samples present pitfalls, making an adequate sample size essential for deriving meaningful conclusions. It is imperative that sample size estimation is conducted *a priori*, before data collection begins, as *post hoc* power calculations are not appropriate in ensuring the reliability and validity of study findings.

References

1. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010. JAMA 2012;307:1838–47.
2. Schuster R, Kaiser T, Terhorst Y, Messner EM, Strohmeier LM, Laireiter AR. Sample size, sample size planning, and the impact of study context: systematic review and recommendations by the example of psychological depression treatment. Psychol Med 2021;51:902–8.
3. Christogiannis C, Nikolakopoulos S, Pandis N, Mavridis D. The self-fulfilling prophecy of post-hoc power calculations. Am J Orthod Dentofacial Orthop 2022;161:315–7.
4. Machin D, Campbell MJ, Tan SB, Tan SH. Sample sizes for clinical, laboratory and epidemiology studies. 4th ed. New York: Wiley-Blackwell; 2018.
5. Roscoe JT. Fundamental research statistics for the behavioral science. International series in decision process. 2nd ed. New York: Holt, Rinehart and Winston; 1975.
6. O'Hara J. How I do it: sample size calculations. Clin Otolaryngol 2008;33:145–9.
7. Staner L. Surrogate outcomes in neurology, psychiatry, and psychopharmacology. Dialogues Clin Neurosci 2006;8:345–52.

Suggested Readings

Barclay DW, Higgins CA, Thompson R. The partial least squares approach to causal modeling: Personal computer adoption and use as illustration. Technology Studies 1995; 2(2): 285–309.

Gorsuch RL. Factor analysis, 2nd ed. New Jersey: Lawrence Erlbaum; 1983.

Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991; 26(3): 499–510.

Hatcher L. A step-by-step approach to using the SAS system for factor analysis and structural equation modeling. North Carolina: SAS Institute; 1994.

Kline RB. Principles and practice of structural equation modeling. New York: The Guilford Press; 2016.

Krejcie RV, Morgan DW. Determining sample size for research activities. Educational and Psychological Measurement 1970;30:607–10.

Lehr R. Sixteen *S*-squared over *D*-squared: A relation for crude sample size estimates. Stat Med 1992;11:1099–102.