Review

Sample size determination: Principles and applications for health research

Bayu Hari Mukti[1,2*]

[1]*Department of Industrial Engeneering, Sari Mulia University, Indonesia*
[2]*Department of Research and Publication, Ananda Health and Education, Indoensia*

[*]Corresponding author: muktibh.works@gmail.com

ARTICLE INFO	ABSTRACT
	Background: Determining an appropriate sample size is a crucial aspect of research design, ensuring validity, reliability, and generalizability of findings. An inadequate sample size increases the risk of Type II errors, while an excessively large sample may lead to resource inefficiencies and a higher likelihood of Type I errors. Understanding the principles of sample size determination, including statistical power, confidence levels, and margin of error, is essential for producing accurate and meaningful research outcomes. *Objective*: This review explores the principles of sample size determination, calculation methods for various research designs, and practical applications. It also discusses challenges in determining the optimal sample size and examines international guidelines, such as those issued by the World Health Organization (WHO), to enhance the accuracy and credibility of research findings. *Discussion*: Sample size determination varies depending on research design, including surveys, experiments, and clinical trials. This review highlights key statistical considerations such as confidence intervals, statistical power, and the role of design effects. Additionally, practical challenges such as resource constraints, parameter misestimation, and population diversity are discussed. Technological advancements, including statistical software, are also examined for their role in improving sample size calculations and research efficiency. *Conclusion*: Adhering to established principles and leveraging modern tools for sample size determination enables researchers to optimize study designs and enhance the validity of findings. Implementing international guidelines minimizes bias and ensures the robustness of results. Ultimately, accurate sample size estimation contributes to high-quality scientific studies that support evidence-based decision-making and progress across various disciplines.

1. Introduction

Determining sample size is a fundamental step in scientific research design, aimed at ensuring the validity, reliability, and generalizability of research findings. Every study, whether in health, social sciences, education, or natural sciences, requires an appropriate sample size to produce accurate conclusions. An overly small sample size can lead to a Type II error, which

is the failure to detect a significant difference when one actually exists. Conversely, an excessively large sample size may result in resource wastage and an increased likelihood of Type I errors, which involve accepting a false hypothesis. A robust sample size calculation is essential for detecting the true effect within a study (Biau et al., 2008; Desachy et al., 2024; Sivasamy, 2023; Vasudevan, 2024).

In research, sample size determination is not solely a mathematical exercise but also involves aligning with the study's objectives. For instance, in survey research, the sample size is calculated based on the number of respondents needed to obtain results that accurately represent the population while controlling for a specified margin of error. Conversely, in experimental or clinical studies, determining sample size is more complex, as it must consider the effectiveness of the intervention, along with associated risks and benefits (Charan & Biswas, 2013).

Statistical power, a key concept in sample size determination, refers to the probability of detecting a true difference if it exists. This power is heavily dependent on the sample size used. An appropriate sample size is critical for optimizing the likelihood of detecting significant findings in research. For example, in clinical trials testing a new drug, a large sample size is often required to ensure that findings on the drug's effectiveness are not due to chance or random fluctuations (Cohen, 2013; Patel, 2024). Various methods are used to calculate sample size, and these methods can vary depending on the research design. In survey research, sample size calculations typically rely on estimates of population variability and the desired confidence level. In clinical experiments, however, determining sample size requires additional considerations, including expected intervention effectiveness and anticipated variability. This process demands a deep understanding of statistical theories and proper contextualization aligned with the study's type and purpose.

Beyond theoretical calculations, sample size determination also involves practical and logistical considerations. Researchers must balance the desire for a sufficiently large sample size with the constraints of available resources. In some cases, determining the sample size can pose significant challenges. A sample that is too small may reduce the study's ability to detect valid results, while an excessively large sample may increase unnecessary costs and time. Additionally, population diversity and challenges in data collection can influence how sample size is calculated.

International organizations, such as the World Health Organization (WHO), have

provided comprehensive guidelines for determining sample sizes in various types of research of health. For instance, WHO has established specific protocols for clinical studies and epidemiological surveys, offering standards to ensure that research outcomes are valid and widely applicable (WHO, 2008). Understanding and adhering to these guidelines is crucial for researchers to ensure the accountability and scientific contribution of their studies.

This article aims to explore the fundamental principles of sample size determination, the calculation methods used across different research designs, and its applications and challenges in various fields. It will present approaches relevant to sample size determination for survey, experimental, and clinical research. Key statistical concepts, including statistical power, margin of error, and confidence level, will serve as the foundation for understanding how sample size influences research outcomes. The article will further discuss the application of sample size determination in health, social, and educational research, emphasizing practical strategies employed by researchers in real-world scenarios. Resource constraints, including time and funding, often influence decisions regarding optimal sample size. Therefore, applying effective sample size determination methods is critical to ensure the smooth conduct of research without compromising validity or accuracy.

Additionally, this article will review international guidelines on sample size determination, such as those issued by WHO and other organizations. These guidelines provide essential references for researchers to ensure that their studies comply with internationally accepted standards, minimizing bias and statistical errors. Furthermore, the discussion will include technological advancements and emerging approaches in sample size determination. With the continuous development of analytical technologies, novel methods for calculating sample size with greater accuracy and efficiency are becoming increasingly available. Ultimately, this article aims to provide researchers with deeper insights into the importance of sample size determination in scientific research and its impact on research outcomes. It will also highlight the challenges associated with determining sample size and offer recommendations on adopting innovative methods to optimize sample size calculations in future studies. By addressing these aspects, the article seeks to contribute to improving the quality and reliability of research across diverse disciplines.

2. Fundamental Principles of Sample Size Determination

Determining sample size is grounded in several statistical principles aimed at ensuring that the research sample accurately represents the target population. These principles include

the key research parameters, confidence level, margin of error, statistical power, and design effect. Together, these principles are interrelated and function to ensure that the chosen sample is truly representative of the broader population. Proper sample size determination minimizes bias, enhances estimation accuracy, and ensures broader applicability of the research findings.

Key research parameters

In the domain of health research, key research parameters refer to the critical variables that underpin the study's design and analytical framework. These parameters commonly include prevalence, incidence rates, means, proportions, and variances, each playing a pivotal role in shaping the methodological rigor of a study. Among these, prevalence—defined as the proportion of individuals in a population who exhibit a particular condition at a specified time—is frequently employed as a foundational parameter for calculating sample size, especially in epidemiological studies.

The relationship between prevalence and sample size exhibits an inverse proportionality, which has significant implications for study design:

- High prevalence: When the prevalence of a condition is high within the target population, a smaller proportion sample size is sufficient to achieve a desired level of confidence and precision. This is because the larger number of cases in the population increases the likelihood of capturing sufficient variability and reduces the margin of error in the estimates.

- Low prevalence: Conversely, conditions characterized by low prevalence necessitate a larger proportion sample size to attain comparable levels of accuracy and reliability. This is particularly critical for rare diseases, where the sparse occurrence of cases amplifies the potential for sampling error and threatens the robustness of the estimations.

In studies of rare diseases, although the available population may be limited, a larger sample size is often necessary to ensure that the findings are reliable and generalizable (Moser & Kalton, 2017). For example, while investigating conditions with low prevalence rates, researchers need to collect more data to achieve robust estimations and minimize uncertainty (Lohr, 2019).

Confidence Level

The confidence level represents the probability that the sample estimate falls within a specific range around the true population parameter. It reflects how certain researchers can

be that the interval includes the actual population value. Commonly used confidence levels in research are 95% and 99%. A 95% confidence level implies a 95% probability that the sample estimate lies within the calculated range of the true population parameter. A 99% confidence level provides even greater certainty, increasing the probability to 99%.

These confidence levels are directly associated with z-scores (Daniel & Cross, 2013), which are used in sample size calculations:

− 95% confidence level: z-score = 1.96

− 99% confidence level: z-score = 2.58.

Higher confidence levels necessitate larger sample sizes to ensure narrower confidence intervals, thereby enhancing the precision and reliability of statistical estimates. Specifically, a 99% confidence level results in a more stringent interval, requiring a larger sample size to achieve the desired level of accuracy (Maxwell et al., 2008). Conversely, a 95% confidence level, while slightly broader, allows for a more manageable sample size without significantly compromising statistical rigor.

The selection of an appropriate confidence level is contingent upon the research context. In medical and clinical research, where decision-making has direct implications for patient outcomes and treatment efficacy, a 99% confidence level is often preferred to minimize the risk of Type I and Type II errors (Celentano & Szklo, 2018). In contrast, social science research and large-scale surveys frequently adopt a 95% confidence level, as it provides a practical balance between precision and feasibility, considering constraints such as cost, time, and respondent burden (Lohr, 2021).

Margin of Error

The margin of error represents the acceptable range of deviation between the sample estimate and the true population parameter. It quantifies the degree of uncertainty that researchers are willing to tolerate in their results. For example, in public opinion surveys, the margin of error is often used to describe how far the sample estimates may deviate from the actual population parameter.

A smaller margin of error necessitates a larger sample size, as researchers seek to ensure that statistical estimates closely approximate the true population parameters. This inverse relationship between margin of error and sample size underscores the critical balance between statistical precision and the logistical constraints of data collection, such as time, cost, and respondent accessibility (Kelley et al., 2023). A well-calibrated margin of error is

essential for maintaining the validity and reliability of research findings while ensuring methodological feasibility.

The selection of an appropriate margin of error is highly context-dependent. In social research, a margin of error of approximately 5% is often considered acceptable, reflecting a reasonable trade-off between precision and practicality (Lohr, 2019). This level of tolerance is particularly common in opinion polling and large-scale surveys, where minor deviations are unlikely to significantly alter the interpretation of aggregate trends.

Conversely, in medical and clinical research, where decision-making directly influences patient outcomes, considerably smaller margins of error—typically 1% or 2%—are required to achieve a high degree of accuracy (Wang & Ji, 2020). For instance, in randomized controlled trials evaluating the efficacy of novel therapeutic interventions, stringent margins of error minimize the risk of Type I (false positive) and Type II (false negative) errors, thereby ensuring that clinical conclusions are both statistically robust and clinically meaningful (Shreffler & Huecker, 2025). In contrast, exploratory or preliminary studies may tolerate broader margins, allowing researchers to identify patterns and generate hypotheses without the constraints of large sample sizes (Conroy, n.d.).

Statistical power

Statistical power refers to the probability that a study will detect a true effect if one exists. A sufficiently high statistical power ensures that the study can identify significant differences when they truly occur. Commonly, a minimum acceptable statistical power is 80%, meaning there is an 80% likelihood of correctly identifying a significant difference if it exists (Cohen, 2013; Schober & Vetter, 2019).

Statistical power depends on three key factors:

− Sample Size: Larger sample sizes enhance statistical power by reducing random error and increasing the precision of parameter estimates. As variability decreases, the likelihood of detecting true effects improves (Lakens, 2022).

− Effect Size: The magnitude of the relationship or difference being measured plays a crucial role in determining power. Larger effect sizes require smaller sample sizes to reach statistical significance, whereas smaller effect sizes demand larger sample sizes to achieve the same level of power (Sullivan & Feinn, 2012).

− Significance Level (α): The chosen α level (e.g., 0.01 or 0.05) directly affects power. Lower α levels impose stricter thresholds for significance, thereby reducing power and

increasing the risk of Type II errors—failing to detect a true effect (Biau et al., 2008; Faul et al., 2007).

In experimental research, inadequate power due to small sample sizes can lead to misleading conclusions, particularly by increasing the likelihood of false negatives. A power level of at least 80% is typically recommended to balance Type I and Type II error risks, ensuring that findings are both statistically robust and practically meaningful (Kraemer & Blasey, 2016). Thus, meticulous sample size determination is essential for generating reliable, valid, and generalizable results in scientific inquiry.

Design Effect

The design effect accounts for variability introduced by sampling techniques that are not entirely random, such as cluster sampling or stratified sampling. In cluster sampling, for instance, individuals within the same cluster tend to share similar characteristics, increasing variability in the estimates and reducing accuracy.

To address the increased variability, researchers apply a design effect coefficient to adjust the required sample size. The formula for effective sample size adjustment is:

$$n_{adjusted} = n \ x \ DE \qquad\qquad (1)$$

Where:

$n_{adjusted}$ = adjusted sample size

n   = initial calculated sample size

DE   = design effect coefficient (2 is most used if intra-cluster correlation coefficien (ρ) is unknown

Failing to account for the design effect can lead to underestimating the required sample size, resulting in biased or unrepresentative findings. Proper adjustments ensure that the study's results remain valid despite the complexities introduced by non-random sampling designs (Cochran, 1977). By incorporating both statistical power and design effect into sample size calculations, researchers enhance the robustness and applicability of their study findings.

3. Application of Sample Size Determination

Sample size determination is an essential component in research design across various fields, including public health, clinical research, education, social studies, and laboratory experiments. Each discipline employs tailored approaches based on the specific characteristics and objectives of the study, both methodologically and contextually. A comprehensive understanding of sample size application in these fields is critical to ensure the reliability of

research outcomes and accurate generalization of findings.

Public health

In public health research, sample size determination heavily relies on the prevalence of the health issue under investigation. These studies often employ sampling techniques that provide a representative population while efficiently managing resource limitations.

- Sample determination based on prevalence

In studies involving the prevalence of specific diseases, such as type 2 diabetes, the sample size is calculated based on known initial prevalence, margin of error, and confidence level. This approach ensures that population variability is accounted for, allowing for generalizable results. The standard formula for sample size calculation is applied as follows:

- Unlimited population $\quad n = \dfrac{z^2 \times \hat{p}(1-\hat{p})}{\varepsilon^2}$ (2)

- Finite population $\quad n' = \dfrac{n}{1+\dfrac{z^2 \times \hat{p}(1-\hat{p})}{\varepsilon^2 N}}$ (3)

where

$\quad$ z is the z score

$\quad$ ε is the margin of error

$\quad$ N is the population size

$\quad$ p̂ is the population proportion

Case example:

In an urban area study, the initial prevalence of type 2 diabetes is estimated at 20%. Researchers aim to estimate the prevalence with a margin of error of ±5% and a 95% confidence level. Using the formula for an infinite population, the sample size is calculated as follows:

$$n = \frac{1.96^2 \cdot 0.2 \cdot (1 - 0.2)}{0.05^2}$$

$$n = \frac{3.8416 \cdot 0.2 \cdot 0.8}{0.0025}$$

$$n = \frac{0.6147}{0.0025} = 246$$

This means that researchers need a minimum of 246 respondents to ensure the study results can be generalized to the population. Table 1 presents the minimum sample size required based on the prevalence of the disease. For unknown prevalence, a value of 50% can be used as a conservative estimate. For further details, refer to the WHO Sample Size Determination for Health Studies (Lwanga et al., 1991).

Table 1. Minimum sample size required based on the prevalence of the disease

Prevalence (%)	Margin of Error (%)	Minimum sample size	
		Confidence Level 95%	Confidence Level 99%
5	5	73	127
	10	19	32
10	5	139	240
	10	35	60
15	5	196	340
	10	49	85
20	5	246	427
	10	62	107
25	5	289	500
	10	73	125
30	5	323	560
	10	81	140
50	5	385	666
	10	97	167

- Sample size base on prevalence in stratified sampling technique

  To study the prevalence of malnutrition in children across regions with significant socio-economic variations, the stratified random sampling technique is often utilized.

  Case example:

  The study is conducted in three regions (urban, rural, and remote) with varying malnutrition prevalence: 5%, 15%, and 25%, respectively. The researcher divides the population into strata based on region and then randomly selects samples from each stratum. The sample size for each stratum is determined based on the prevalence in that region, resulting in a sufficiently large total sample to achieve a 95% confidence level. This strategy helps identify regional variations and provides a more accurate picture of the malnutrition issue (Cochran, 1977).

- Sample size base on case population

  A representative sample size is a crucial aspect of quantitative research to ensure the validity and reliability of results. One commonly used method in survey and epidemiological research is Slovin's formula, which allows researchers to determine the required number of respondents based on a known population while considering the margin of error.

$$n = \frac{N}{1 + Ne^2} \qquad (4)$$

Where:

n is required sample size

N is total population

e is margin of error

This method is frequently used in research with limited resources and time, enabling researchers to determine a sufficiently representative sample size without conducting a full census. Advantages of Slovin's Formula:

- Simplicity and ease of use: The formula provides a straightforward method for calculating sample size, making it accessible to researchers without advanced statistical backgrounds.

- Applicability with limited information: Slovin's formula is particularly useful when little is known about the population's behavior or variability, offering a starting point for sample size determination in the absence of detailed data.

- Flexibility across disciplines: The formula is utilized in various fields, including market research, social sciences, and quality control, to estimate sample sizes that balance accuracy and practicality.

Disadvantages of Slovin's Formula:

- Assumption of random sampling: The formula presumes that the sample is drawn using simple random sampling. In populations with distinct subgroups, applying Slovin's formula without adjustments may lead to unrepresentative samples.

- Lack of consideration for population variability: Slovin's formula does not account for the inherent variability within the population. In cases where the population has high variability, the formula may underestimate the required sample size, compromising the accuracy of results.

- Potential for misuse: Due to its simplicity, there is a risk of misapplying Slovin's formula in complex research scenarios where more sophisticated sampling techniques are warranted. Researchers must ensure the formula's assumptions align with their study design to avoid erroneous conclusions.

Case example:

A researcher aims to assess the prevalence of hypertension in the adult population of City X, which has a total population of 50,000 people. To ensure research efficiency, a 5% (0.05) margin of error is chosen, as commonly accepted in social and epidemiological studies.

Sample size determination (Mukti, 2025)

Using Slovin's formula:

$$n = \frac{50000}{1 + 50000\,(0.05)^2}$$

$$n = \frac{50000}{1 + 125}$$

$$n = \frac{50000}{126}$$

$$n = 397$$

Thus, a minimum of 397 respondents is required to ensure the study results are sufficiently representative and statistically valid.

- Power analysis in clinical research

Determining sample size in clinical research is a crucial step as it directly impacts the validity, reliability, and safety of the research findings. Unlike survey or social research, clinical studies require additional considerations related to the risks of interventions, expected benefits, and potential side effects. Therefore, sample size determination typically uses more complex statistical methods to ensure accurate results and generalizability to a larger population (Friedman et al., 2015).

Power analysis is one method commonly used to determine sample size. It incorporates factors such as the desired significance level (typically 0.05) and a statistical power of at least 80% to detect a significant difference between the treatment and control groups (Sullivan & Feinn, 2012).

This method takes into account several key factors:

- Significance Level (α): The probability of incorrectly accepting a statistically significant result, typically set at 5% (α = 0.05).

- Statistical Power: The probability of detecting a significant difference if one truly exists, with a minimum value of 80% (1−β = 0.84).

- Desired Clinical Effect (Effect Size): The magnitude of the expected difference between the intervention and control groups.

- Data Variability: The level of spread or heterogeneity in the data within the population.

The general formula is:

$$n = \frac{(Z_\alpha + Z_\beta)^2 \cdot 2 \cdot \sigma^2}{d^2} \tag{5}$$

Where:

Z_α: Z-score for significance level

Z_β: Z-score for statistical power

σ: Standard deviation

d: Expected difference (Mean1 – Mean2) $(Mean_1 - Mean_2)$ (Kim, 2016)

Case example:

In a clinical trial to evaluate the effectiveness of a new drug for hypertension, researchers estimate that the drug will lower the average blood pressure by 10 mmHg compared to the control group, with a standard deviation of 15 mmHg. With a significance level of 5% and statistical power of 80%, the sample size for each group is calculated as:

$$n = \frac{(Z_\alpha + Z_\beta)^2 \cdot 2 \cdot \sigma^2}{d^2}$$

$$n = \frac{(1.96 + 0.84)^2 \cdot 2 \cdot 15^2}{10^2}$$

$$n = \frac{7.84 \cdot 450}{100} = \frac{3528}{100} = 35.28$$

Thus, the minimum sample size required per group is 36 subjects.

- 30 x 7 Cluster Sampling Technique

The 30 x 7 cluster sampling technique is employed to ensure efficiency in obtaining a representative sample. This method is commonly used in immunization coverage assessments, alongside Lot Quality Assessment Sampling (LQAS) (Hoshaw-Woodard, 2001). The technique involves randomly selecting 30 clusters, with each cluster consisting of 7 respondents representing the group.

Case example:

A study to evaluate the measles immunization coverage among children aged 9 months to 5 years in a rural area. From the target population, the researcher randomly selects 30 villages as clusters. In each village, 7 children are identified using systematic sampling. This method enables faster data collection by reducing the number of locations that need to be accessed, which is particularly beneficial in areas with accessibility challenges. Additionally, studies have shown that alternative sampling designs can increase efficiency by up to 1.7 times compared to Simple Random Sampling (SRS), allowing researchers to gather more relevant data in a shorter amount of time (Parsaeian et al., 2021; Yin et al., 2017).

Experimental studies

In laboratory experiments, sample size determination is based on considerations of data variability and the desired level of significance. Research testing the effectiveness of a new drug or chemical substance on animal models, for example, requires careful sample size calculation to ensure that the results are statistically significant and can be generalized to the human population. Furthermore, experiments involving animal models often account for the effects of external variables such as diet or environmental conditions, which can influence the experimental outcomes. Therefore, researchers must take these factors into account when determining sample size to minimize bias and unwanted variability. By calculating an appropriate sample size, experimental research can ensure that the conclusions drawn truly reflect the intended effects and are scientifically accountable (Cohen, 2013).

Apart from variability and significance level, sample size in laboratory experiments is also influenced by the practical importance of the outcomes being sought. For example, in biomedical research testing the therapeutic effects of a compound, a larger sample size may be required to detect very small but clinically significant differences that might not be detectable in experiments with smaller sample sizes (Biau et al., 2008).

The application of sample size determination varies greatly across different research fields. In public health, clinical research, education, social studies, and laboratory experiments, sample size plays a critical role in ensuring the validity and reliability of research outcomes. By understanding the basic principles and specific approaches applied in each field, researchers can design studies that are more effective, efficient, and representative, ensuring that research findings are broadly applicable and useful for decision-making in their respective fields.

4. Utilization of Software and Web-Based Applications in Sample Size Determination

Technological advancements have significantly improved the efficiency and accuracy of statistical analyses in health research. One critical area benefiting from these advancements is sample size determination, a fundamental component of research methodology that ensures validity, reliability, and generalizability of findings (Patel, 2024). Traditional manual calculations for determining sample size are time-consuming and prone to errors, leading to an increasing reliance on software and web-based applications for more precise and efficient estimations (Lakens, 2022). These digital tools integrate complex statistical models, reducing computational burdens and facilitating robust study designs across diverse research fields,

including epidemiology, clinical trials, and public health studies (Wang & Ji, 2020).

Several statistical software and web-based applications have been developed to assist researchers in determining sample size. These tools utilize key statistical parameters such as statistical power, effect size, confidence intervals, and design effect to ensure rigorous study design. Notable software includes:

1. Calculator.net Sample Size Calculator: A web-based tool incorporating WHO's sample size determination guidelines, providing straightforward calculations for different study designs.

2. Slovin's Formula Calculator (Statology.org): A convenient web-based tool allowing researchers to quickly determine sample size for population studies using Slovin's formula, useful in survey research and epidemiological studies.

3. G*Power: A widely used, open-source program that calculates sample size for various statistical tests, including t-tests, ANOVA, and regression models. G*Power is particularly useful for behavioral and biomedical research, providing researchers with flexibility in adjusting power and effect size to optimize sample size estimation.

4. PASS (Power Analysis and Sample Size Software): A commercial software offering advanced features for precise sample size determination across multiple study designs, including survival analysis and nonparametric tests. PASS is frequently utilized in medical and pharmaceutical research due to its comprehensive statistical capabilities.

5. Epi Info: Developed by the Centers for Disease Control and Prevention (CDC), this free software provides researchers with epidemiological tools, including sample size calculators for cross-sectional, cohort, and case-control studies. Epi Info is particularly beneficial for public health researchers working in resource-limited settings.

6. OpenEpi: A web-based application offering similar functionalities to Epi Info, allowing users to conduct power and sample size calculations for epidemiological research. OpenEpi is advantageous due to its accessibility and ease of use, requiring no installation.

7. nQuery: A premium software commonly used in clinical trials, offering advanced statistical modeling for determining optimal sample sizes in randomized controlled trials (RCTs). nQuery supports Bayesian and frequentist approaches, making it suitable for modern clinical research methodologies.

Case Example

To illustrate the significance of software-based sample size determination, consider a

clinical trial evaluating the effectiveness of a new antihypertensive drug. The study aims to detect a mean difference of 10 mmHg in systolic blood pressure between the treatment and control groups, with an expected standard deviation of 15 mmHg. Using G*Power, the researcher sets the parameters as follows:

− Statistical test: Independent samples t-test

− Effect size (d): 0.67 (calculated as mean difference divided by standard deviation)

− Power (1-β): 0.80

− Significance level (α): 0.05

The output from G*Power suggests a required sample size of 36 participants per group to achieve adequate statistical power. By utilizing software, the researcher efficiently determines an appropriate sample size without manual calculations, minimizing the risk of underpowered or overpowered study designs.

the growing emphasis on large-scale and remote studies, web-based sample size calculators have become indispensable. Platforms like OpenEpi allow researchers to calculate sample sizes for disease prevalence studies conducted in remote populations with limited statistical expertise. Additionally, Sample Size Calculator offers a simplified approach for health researchers adhering to WHO's sample size determination methods. Meanwhile, Slovin's Formula Calculator enables quick and effective sample size computation for survey-based studies, particularly useful for researchers conducting large-scale epidemiological research.

Additionally, cloud-based applications integrated with machine learning algorithms are emerging as innovative solutions in sample size estimation. These platforms analyze historical datasets and real-time epidemiological trends to refine sample size calculations dynamically. This approach is particularly valuable in pandemic-related research, where sample size requirements may fluctuate due to evolving transmission rates and population heterogeneity.

The adoption of software and web-based tools for sample size calculation presents numerous advantages, including:

− Efficiency: Automated calculations reduce time and effort compared to manual computations, enabling researchers to focus on study design and data interpretation.

− Accuracy: Statistical software minimizes computational errors, ensuring robust and reproducible sample size estimates.

− Accessibility: Open-source and web-based applications facilitate sample size

determination for researchers in resource-constrained settings.

However, challenges remain in the widespread adoption of these tools. Some software requires advanced statistical knowledge, making them less accessible to researchers without formal training in biostatistics (Weissgerber et al., 2016). Additionally, commercial software solutions may be cost-prohibitive for researchers in low-income settings, limiting their applicability in global health studies.

As technology continues to evolve, integrating artificial intelligence (AI) and machine learning into sample size determination is expected to enhance precision and adaptability. Future research should explore AI-driven predictive models that optimize sample size based on real-time data trends and contextual factors. Moreover, increasing accessibility to open-source software will bridge the gap between resource-rich and resource-limited researchers, promoting equity in scientific investigations.

5. Conclusion

Determining an appropriate sample size is a fundamental aspect of research design that directly impacts the validity, reliability, and applicability of findings. This review highlights the importance of understanding the underlying principles, including statistical power, confidence levels, and margins of error, which guide the calculation of sample sizes across diverse research designs. By adhering to international guidelines and leveraging modern technologies, researchers can achieve more accurate and efficient outcomes. Ultimately, optimizing sample size determination enhances the quality of scientific studies, ensuring their contribution to evidence-based decision-making and broader scientific advancements.

6. Conflict of interest

All authors declare no conflict of interest.

7. References

Biau, D. J., Kernéis, S., & Porcher, R. (2008). Statistics in Brief: The Importance of Sample Size in the Planning and Interpretation of Medical Research. *Clinical Orthopaedics and Related Research*, *466*(9), 2282–2288. https://doi.org/10.1007/s11999-008-0346-9

Celentano, D. D., & Szklo, M. (2018). *Gordis Epidemiology* (6th ed.). Elsevier.

Charan, J., & Biswas, T. (2013). How to Calculate Sample Size for Different Study Designs in Medical Research? *Indian Journal of Psychological Medicine*, *35*(2), 121–126. https://doi.org/10.4103/0253-7176.116232

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

Conroy, R. (n.d.). *Sample Size - A rough Guide*. NDI. Retrieved February 15, 2025, from https://www.ndi.org/sites/default/files/samplesizecalculation.pdf?utm_source=chatgpt.com

Daniel, W. W., & Cross, C. L. (2013). *Biostatistics: A Foundation for Analysis in the Health Sciences* (10th Edition). John Wiley & Sons.

Desachy, T., Thevenet, M., Garcia, S., Lightning, A., Didier, A., Mandairon, N., & Kuczewski, N. (2024). *Enhancing Statistical Power While Maintaining Small Sample Sizes in Behavioral Neuroscience Experiments Evaluating Success Rates*. https://doi.org/10.1101/2024.07.25.605060

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of Clinical Trials*. Springer International Publishing. https://doi.org/10.1007/978-3-319-18539-2

Hoshaw-Woodard, S. (2001). *Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage*. Department of Vaccones and Biologicals - WHO.

Kim, H.-Y. (2016). Statistical notes for clinical researchers: Sample size calculation 1. comparison of two independent sample means. *Restorative Dentistry & Endodontics*, *41*(1), 74. https://doi.org/10.5395/rde.2016.41.1.74

Kraemer, H. C., & Blasey, C. (2016). *How Many Subjects?: Statistical Power Analysis in Research*. SAGE Publications, Ltd. https://doi.org/10.4135/9781483398761

Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.33267

Lohr, S. L. (2019). *Sampling*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429296284

Lohr, S. L. (2021). *Sampling - Design and Analysis*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429298899

Lwanga, S. K., Lemeshow, S., & World Health Organization. (1991). *Sample size determination in health studies: a practical manual*. World Health Organization.

Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample Size Planning for Statistical Power and Accuracy in Parameter Estimation. *Annual Review of Psychology*, *59*(1), 537–563. https://doi.org/10.1146/annurev.psych.59.103006.093735

Moser, C. A., & Kalton, G. (2017). *Survey Methods in Social Investigation*. Routledge. https://doi.org/10.4324/9781315241999

Parsaeian, M., Mahdavi, M., Saadati, M., Mehdipour, P., Sheidaei, A., Khatibzadeh, S., Farzadfar, F., & Shahraz, S. (2021). Introducing an efficient sampling method for national surveys with limited sample sizes: application to a national study to determine quality and cost of healthcare. *BMC Public Health*, *21*(1), 1414. https://doi.org/10.1186/s12889-021-11441-0

Patel, D. (2024). Sample Size Estimation in Clinical Trials. *National Journal of Community Medicine*, *15*(06), 503–508. https://doi.org/10.55489/njcm.150620243815

Schober, P., & Vetter, T. R. (2019). Sample Size and Power in Clinical Research. *Anesthesia & Analgesia*, *129*(2), 323–323. https://doi.org/10.1213/ANE.0000000000004316

Shreffler, J., & Huecker, M. R. (2025). *Type I and Type II Errors and Statistical Power*.

Sivasamy, S. (2023). Sample size considerations in research. *Endodontology*, *35*(4), 304–308. https://doi.org/10.4103/endo.endo_235_23

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the *P* Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Vasudevan, S. (2024). Sample Size Calculation in Various Medical Research. *International Journal of Medical Sciences and Nursing Research*, *4*(3), 22–29. https://doi.org/10.55349/ijmsnr.2024432229

Wang, X., & Ji, X. (2020). Sample Size Estimation in Clinical Research. *Chest*, *158*(1), S12–S20. https://doi.org/10.1016/j.chest.2020.03.010

Weissgerber, T. L., Garovic, V. D., Milin-Lazovic, J. S., Winham, S. J., Obradovic, Z., Trzeciakowski, J. P., & Milic, N. M. (2016). Reinventing Biostatistics Education for Basic Scientists. *PLOS Biology*, *14*(4), e1002430. https://doi.org/10.1371/journal.pbio.1002430

WHO. (2008). *WHO child growth standards : Training course on child growth assessment.* WHO.

Yin, G., Li, A., & Verger, A. (2017). Spatiotemporally Representative and Cost-Efficient Sampling Design for Validation Activities in Wanglang Experimental Site. *Remote Sensing*, *9*(12), 1217. https://doi.org/10.3390/rs9121217