Acta Medica Alanya

EDITORIAL

EDİTÖRYAL

Acta Medica Alanya

2025;9(3): 162-164

DOI: 10.30565/medalanya.1850488

The Importance of Sample Size Calculation and Power Analysis in Research Studies

Araştırma Çalışmalarında Örneklem Büyüklüğü Hesaplamasının ve Güç Analizinin Önemi

Özkan Köse[1], Ahmet Aslan[2]*

1.University of Health Sciences, Antalya Education and Research Hospital, Department of Orthopedics and Traumatology, Antalya, Türkiye,
2.Medical School of Alanya Alaaddin Keykubat University, Department of Orthopedics and Traumatology, Alanya/Antalya, Türkiye

ABSTRACT

Decisions about sample size and statistical power are often treated as a technical afterthought in research proposals, squeezed into a single sentence or delegated entirely to software. Yet these decisions sit at the intersection of scientific validity, feasibility, and ethics and directly determine whether a study can answer its primary question in a trustworthy way. Sample size calculation and power analysis should no longer be regarded as optional technicalities appended at the end of a protocol. They are central to scientific integrity, determining whether a study can credibly answer its research question, justify the burdens placed on participants, and make a meaningful contribution to the literature. In this editorial, we aim to draw attention to sample size calculation and power analysis, outlining their conceptual foundations, practical implementation, and ethical implications in contemporary research.

Key Words: Sample Size, Power Analysis, Research Studies

ÖZ

Araştırma önerilerinde örneklem büyüklüğü ve istatistiksel güç hakkındaki kararlar genellikle teknik bir sonradan düşünme unsuru olarak ele alınır, tek bir cümleye sıkıştırılır veya tamamen yazılıma devredilir. Oysa bu kararlar bilimsel geçerlilik, uygulanabilirlik ve etik konularının kesiştiği noktada yer alır ve bir çalışmanın temel sorusunu güvenilir bir şekilde yanıtlayıp yanıtlayamayacağını doğrudan belirler. Örneklem büyüklüğü hesaplaması ve güç analizi artık bir protokolün sonuna eklenen isteğe bağlı teknik ayrıntılar olarak görülmemelidir. Bunlar bilimsel bütünlüğün merkezinde yer alır ve bir çalışmanın araştırma sorusunu güvenilir bir şekilde yanıtlayıp yanıtlayamayacağını, katılımcılar üzerindeki yükleri haklı çıkarıp çıkaramayacağını ve literatüre anlamlı bir katkı sağlayıp sağlayamayacağını belirler. Bu editöryal yazıda, örneklem büyüklüğü hesaplaması ve güç analizine dikkat çekmeyi, kavramsal temellerini, pratik uygulamalarını ve çağdaş araştırmalardaki etik sonuçlarını özetlemeyi amaçlıyoruz.

Anahtar Kelimeler: Örneklem Büyüklüğü, Güç Analizi, Araştırma Çalışmaları

Received:27.12.2025 Accepted:29.12.2025 Published (Online):31.12.2025

* Corresponding Author: Ahmet Aslan, Alanya Alaaddin Keykubat University, Faculty of Medicine, Department of Orthopedics and Tarumatology, Alanya/Antalya, Türkiye. Phone: 05056462411 / mail: ahmet.aslan@alanya.edu.tr

ORCID: 0000-0001-5797-1287

To cited: Köse Ö, Aslan A. The Importance of Sample Size Calculation and Power Analysis in Research Studies. Acta Med. Alanya 2025;9(3): 162-164 DOI: 10.30565/medalanya.1850488

Decisions about sample size and statistical power are often treated as a technical afterthought in research proposals, squeezed into a single sentence or delegated entirely to software. Yet these decisions sit at the intersection of scientific validity, feasibility, and ethics and directly determine whether a study can answer its primary question in a trustworthy way [1–3]. In quantitative research, where we aim to draw inferences from samples to populations, inadequate planning of sample size and power can turn a well-intentioned project into a research-wasting, ethically questionable exercise. Underpowered studies risk missing clinically meaningful effects; overpowered studies can consume unnecessary resources and expose more participants than needed to potential harm [1,2,4–6]. In this editorial, we aim to draw attention to sample size calculation and power analysis, outlining their conceptual foundations, practical implementation, and ethical implications in contemporary research.

Sample size calculation answers a deceptively simple question: How many participants do we need to detect a prespecified effect with acceptable error rates? The answer depends on several key ingredients, including the primary outcome and its

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License.

scale, the anticipated effect size that is considered clinically or practically important, the expected variability in the data, the desired Type I error rate (α), the acceptable Type II error rate (β) and corresponding power (1−β), and the study design, including allocation ratio, clustering, and number of measurements per participant [2,3,7]. An adequately sized study protects against false reassurance. For example, a clinical trial with too few participants may conclude that a new treatment is "not different" from standard care, even though the study was incapable of detecting the true effect. As Altman pointed out decades ago, such trials are not just scientifically weak; they are ethically problematic because participants accept burdens and risks in studies that had little chance of being informative [1,6]. At the same time, sample size is not merely a matter of "bigger is better." Excessively large samples inflate costs, prolong recruitment, and may expose more participants than necessary. The goal is not the most extensive possible study, but a sufficiently precise one, justified by transparent assumptions and aligned with the primary research question [1,2]. Practical tools, such as dedicated software and online calculators, make sample size estimation accessible. However, these tools are only as good as the assumptions entered. In areas like discrete-choice experiments or complex designs, specialized guidance is now available to support more nuanced calculations [4,7].

Power analysis quantifies the probability that a study will detect a true effect of interest given the chosen sample size, effect size, and significance level. In other words, it is the probability of correctly rejecting a false null hypothesis and avoiding a Type II error [2,3]. Conventionally, many fields accept a power of 80% as a minimum standard, acknowledging a 20% risk of failing to detect the planned effect. In some contexts, such as pivotal clinical trials or confirmatory phase III studies, higher power (e.g., 90%) may be warranted, especially when the stakes for patients and policy are high [2,8]. Crucially, power analysis should be prospective and conducted before data collection. Post hoc or "observed" power, calculated after seeing non-significant results, adds little value and can be misleading. Instead, researchers should interpret non-significant findings in light of the a priori detectable effect size and the planned power, asking, "Was this study even capable of detecting a clinically important difference?" [3,8]. In settings such as public health, surgery, or rehabilitation trials, where interventions are costly and logistically complex, robust power analysis helps ensure that true benefits are not overlooked and that results are worth the investment made by funders, clinicians, and participants [2,5].

Sample size and power are tightly interwoven. For a fixed effect size and significance level, increasing sample size increases power, yielding more precise estimates and narrower confidence intervals. Conversely, when resources are constrained, researchers may need to accept lower power or redefine the minimum effect size of interest [2,7,8]. Modern reporting guidelines explicitly recognize this interdependence. The CONSORT statement requires randomized trials to report how sample size was determined, and recent extensions require authors to explicitly state the target difference used in the calculation [8,9]. Despite this, audits show that many published trials either omit sample-size justification or report it incompletely, limiting the reader's ability to judge whether "no difference" actually means "no important difference." [5]. From a design perspective, thinking jointly about sample size and power encourages a transparent conversation among investigators, statisticians, funders, and ethics committees. This conversation should address the effect size that would justify changing practice, the acceptable level of uncertainty, and whether the planned sample is realistic for recruitment and follow-up.

Although guidelines are clear, several challenges persist. Reliable estimates of effect size and variability are often unavailable, especially in novel areas. As indicated by the extant literature, pilot studies, systematic reviews, and data from related populations can inform assumptions; however, they rarely eliminate uncertainty [3, 4, 7]. The constraints imposed by funding, the pressure to publish promptly, and the challenges in recruitment all contribute to an environment in which investigators are compelled to conduct small-scale studies with optimistic assumptions. Meta-research has demonstrated that prevailing academic incentives can systematically favor underpowered research, contributing to irreproducible science and wasted resources [5,10]. In the context of cluster-randomized trials, adaptive designs, survival analyses, or studies with

multiple primary outcomes, the need for meticulous calculations and vigilant oversight of Type I error becomes paramount. Conventional wisdom and rudimentary two-group formulas are frequently insufficient [5,7,8]. Ethicists and methodologists have increasingly argued that decisions regarding sample size must be grounded not only in statistics but also in ethical principles, including respect for participants, fair allocation of risk, and avoidance of futile or needlessly large trials [1,6]. Despite the execution of suitable calculations, the reporting of these results may be deficient or even absent. Journals that endorse CONSORT and analogous guidelines can facilitate this process by requiring clear descriptions of assumptions, methods, and any deviations from the original plan [8,9]. Addressing these challenges requires a dual approach: statistical expertise and a cultural shift that prioritizes quality over quantity in research output.

Sample size calculation and power analysis should no longer be regarded as optional technicalities appended at the end of a protocol. They are central to scientific integrity, determining whether a study can credibly answer its research question, justify the burdens placed on participants, and make a meaningful contribution to the literature [1–3,6]. For investigators, this means engaging with sample size and power early, in close collaboration with biostatisticians, and documenting assumptions transparently. For reviewers and editors, it means treating inadequate or opaque justifications as a serious methodological flaw. For funders and ethics committees, it means recognizing that appropriately powered, well-designed studies serve patients and science far more than a multitude of small, inconclusive projects [5,6,10]. Ultimately, rigorous attention to sample size and power is not merely a statistical nicety; it is a moral and scientific obligation. When done well, it transforms research from a gamble into a disciplined, transparent endeavor that genuinely advances knowledge and improves practice.

Conflict of Interest: No conflict of interest was declared by the author.

Funding sources: The author declared that this article received no financial support.

ORCID and Author contribution: **Ö.K. (0000-0002-7679-9635):** Concept, literature search, writing, critical review. **A.A. (0000-0001-5797-1287):** Critical review, editing.

REFERENCES

1. Altman DG. Statistics and ethics in medical research: III How large a sample? Br Med J. 1980 Nov 15;281(6251):1336-8. doi:10.1136/bmj.281.6251.1336.
2. Kang M, Ragan BG, Park JH. Issues in outcomes research: an overview of randomization techniques for clinical trials. J Athl Train. 2008;43(2):215-21. doi:10.4085/1062-6050-43.2.215.
3. Andrade C. Sample Size and its Importance in Research. Indian J Psychol Med. 2020;42(1):102-103. doi: 10.4103/IJPSYM.IJPSYM_504_19.
4. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. Patient. 2015;8(5):373-84. doi: 10.1007/s40271-015-0118-z.
5. Rutterford C, Taljaard M, Dixon S, Copas A, Eldridge S. Reporting and methodological quality of sample size calculations in cluster randomized trials could be improved: a review. J Clin Epidemiol. 2015;68(6):716-23. doi: 10.1016/j.jclinepi.2014.10.006.
6. Cesana BM, Antonelli P. Sample size calculations in clinical research should also be based on ethical principles. Trials. 2016;17(1):149. doi:10.1186/s13063-016-1277-5.
7. Charan J, Biswas T. How to calculate sample size for different study designs in medical research? Indian J Psychol Med. 2013;35(2):121-6. doi:10.4103/0253-7176.116232.
8. Hopewell S, Chan AW, Collins GS et al. CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials. BMJ. 2025;389:e081124. doi: 10.1136/bmj-2024-081124.
9. Cuschieri S. The CONSORT statement. Saudi J Anaesth. 2019;13(Suppl 1):S27-S30. doi:10.4103/sja.SJA_559_18.
10. Higginson AD, Munafò MR. Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. PLoS Biol. 2016;14(11):e2000995. doi: 10.1371/journal.pbio.2000995.