CME

How to Deal with Sample Size Exploration and its reporting in Clinical Research?

Sada N. Dwivedi

The minimum sample size required to answer a specific research question is unavoidable regardless of study designs, more pertinent in case of randomized controlled trials. Sample size exploration is not required in many situations, including a pilot study; it is simply decided as feasible or a rule of thumb. However, in situations, including a conclusive study like the third and final phase of drug development, it becomes mandatory to ensure optimal power as well as level of confidence to the study. The present write-up aims to address various issues related to sample size considerations for clinical studies in a non-statistical language, focusing on making it easily understandable by medical researchers.

Access this article online

Website: www.cijmr.com

Keywords: Minimum Sample Size, Power of the study, Level of Confidence, Clinical Research

10.58999/cijmr.2022.131

Background

Having spent more than three decades at premier public medical institute of India, All India Institute of Medical Sciences, New Delhi, and working at a reputed private medical college presently, made me realize the need to highlight the tremendous potential of few of the important biostatistical methods and approaches used in carrying out clinical research. Out of numerous steps under clinical research methodology, none is free from a biostatistical feel. Under a study, after the identification of the research question and related hypothesis and objectives of the study, tools of data collection need to be either developed or modified to suite the local need. While doing this, scales of measurements of each variable has its own role. For example, change in a single variable's measurement scale might change the results under epidemiological modeling.[1] In other words, not only changing research questions/hypotheses/objectives need different study designs and analytical approaches but changing scales of measurements of a variable, (including exposure and outcome variables) individually as well as in the combination of other variables may also need varying study designs, analytical methods, and their interpretations. That is why each variable's measurement scale needs to be finalized objectively. Further, out of various statistical components involved in planning, execution, analysis and interpretation of analytical results, sample size exploration is often misunderstood, and attempts are often made to claim coverage of explored minimum sample size. As a result, often related efforts while exploring the required minimum sample size for a specific objective under a study become totally unscientific, leading to distorted results with a likelihood of disastrous implications. This write-up aims to address a few issues related to sample size exploration under the study design of a clinical study.

A clinical study might be either observational or interventional. In contrary to an observational study, under an interventional study, exposure (e.g., drug, yoga) is assigned by the researcher. Further, an observational study might be either descriptive or analytical. A descriptive study does not have a comparator group. In contrary to descriptive study, the analytical study involves a comparator. On similar lines, uncontrolled trial means might exist without a comparator under interventional study. An interventional study without random allocation of study participants (patients/persons) in different study arms remains a non-randomized trial, also referred to as semi-experimental.

International Centre for Health Research, RD Gardi Medical College, Ujjain, Madhya Pradesh, India

Correspondence to: Sada N. Dwivedi, International Centre for Health Research, RD Gardi Medical College, Ujjain, Madhya Pradesh, India. E-mail: dwivedi7@hotmail.com

Submitted: 10/10/2022 Revision: 28/10/2022

Accepted: 15/11/2022 Published: 27/12/2022

How to cite this article: Dwivedi S.N. How to Deal with Sample Size Exploration and its reporting in Clinical Research? Central India Journal of Medical Research. 2022;1(3):2-6.

On the other hand, an interventional study with a random allocation of study participants in different study arms becomes a randomized controlled clinical trial. As a matter of fact, in many situations, sample size exploration goes on similar lines regardless of an observational or interventional study. Again, regardless of observational or interventional study, clinical studies may be broadly classified into two categories, a study with a primary objective as estimation/description or a study with a primary objective as a comparative study. As obvious, a comparative study means a study to test a hypothesis. Each of these two groups of studies needs varying inputs while exploring its related required minimum sample size. The explored sample size is a minimum requirement to reply to a research question scientifically valid. As such, in comparison to explored minimum sample size, covering larger sample size is a welcome step to provide more precise results. However, compared to the explored minimum sample size, a comparatively covered smaller sample size may not be an appropriate step.

Studies with Primary Objective of Estimation/Description

From a public health program perspective, one may estimate the average hemoglobin level among pregnant mothers. However, based on a certain threshold of hemoglobin level, another researcher may like to estimate the burden of anemia among pregnant mothers. Without going into relative importance of these two studies, with a focus on sample size exploration, required inputs for sample size calculation will include variance and required absolute precision (that is, clinically acceptable difference from mean hemoglobin) in first case, where as proportion of anemic pregnant mothers and acceptable absolute error (precision) in it in second case. As such, these two objectives might involve two different sample sizes; hence, there is a need to debate the relative importance of these two objectives and decide accordingly. The common input under these two options is normal deviate at the considered level of confidence.

The level of confidence, complementary to level of significance, means the probability or chance to correctly estimate average hemoglobin level/ proportion of anemic mothers based on sample study to those prevailing in study population. The normal deviation value at a higher confidence level would obviously emerge to be higher and push the required minimum sample size higher. Often, to reduce required minimum sample size, the researchers keep reducing confidence level without noticing its distorted implications. A reduction in confidence level results in an increase in significance level which is nothing but accepting a higher chance of error in estimation at the planning level. For instance, instead of considering 95% level of confidence, if we consider 90%, the level of accepting error (i.e., level of significance) increases from 5% earlier to 10% now. A higher error allowed in sample size exploration will also get reflected in the confidence interval (interval estimates) of the estimated average hemoglobin level/proportion of anemic mothers. Hence, there is need to consider optimal level of confidence while exploring the required minimum sample size. As a rule of thumb, 95% level of confidence is an optimal level while exploring required minimum sample size for a study with the primary objective of estimation/description. Such objectives do not involve power while exploring sample size because they aim to get an estimate closure to population mean/ proportion, not otherwise.

A larger minimum sample size will be required in both cases (mean/ proportion), higher the variance. In other words, in case of proportion, at an acceptable fixed absolute error in proportion (in the denominator of sample size formula) and a fixed level of confidence, the highest required minimum sample size will be at 50%. Hence, in case of a fixed level of absolute error and level of confidence, it is advisable to consider a proportion that is closure to 50% among various proportions reported in the literature. If there is no study on a topic, as a rule of thumb, 50% is considered to plan a study. For example, in our country, universal immunization coverage and HIV sentinel surveillance study are planned with this consideration.

It may be worthwhile to caution here that deciding inputs for sample size exploration under the estimation objective requires exhaustive review literature and an understanding of the public health implications of considered inputs. For example, we need to target estimation in the range which might be relevant from public health point of view. In their absence, there might be an underestimation of the required minimum sample size leading to inaccurate/instable estimates, or overestimation leading to wastage of time/money and other resources.

Studies with Primary Objective of Testing Hypothesis/Comparison

Under this objective, one may decide to compare the average hemoglobin level between two pregnant mothers or the proportion of anemic mothers between these

groups. As pointed out above under objective estimation, these two objectives might need different sample sizes. Also, they will have different public health implications. For example, comparatively average hemoglobin levels in two groups might be statistically significantly different along with either of three conditions: (i) both groups have got a normal range of hemoglobin level; (ii) both groups have got an abnormal range of hemoglobin levels; and (iii) one group has got the normal range and another abnormal range of hemoglobin. Under the first condition, both groups do not require additional attention/ care. Under second condition, regardless of statistical significance, each group needs added care. Under third condition, obviously, group identified with abnormal hemoglobin only will need added care. Hence, one must decide objectively in view of its public health implications. For further discussion, let us presume that objective is to compare the proportion of anemic mothers in two groups.

The required inputs after due thought under this objective involve normal deviate value at the considered level of confidence; normal deviate value at considered power of the study; proportion of anemic mothers in the first group and proportion in the second group; and clinically acceptable effect size. In this case, confidence level may be defined as probability/chance of correctly concluding that both groups have comparable proportion/ burden of anemic mothers. On the other hand, power of the study may be defined as probability/chance of correctly concluding that both groups have not got comparable proportions/burden of anemic mothers. As a convention, level of confidence and power of the study are considered as 95% and 80, respectively. But, under randomized trials, it is advisable to increase them further, like 99 and 90%, respectively. Once we increase them, the calculated sample size for each group will move higher. Further, contrary to the study's power, a normal deviate value at a specific level of confidence remains higher in case of a two-sided hypothesis of the study (e.g., superiority trials) than that in case of a one-sided hypothesis (e.g., non-inferiority trials). As such, clinically acceptable effect size, in this case, needs to be debated and decided in consultation with public health experts. As obvious, being in the denominator, it is inversely correlated with required minimum sample size. In contrast, being in the numerator, the higher the pooled variance and the higher the required minimum sample size. A thorough understanding of these implications might help in objectively exploring the minimum sample size.

Study Design Specific Sample Size Exploration

The details described above mainly relate to the cross-sectional study. However, suppose described principles are understood and conceptualized. In that case, it is easier to use same principles while exploring sample size in case of case-control studies, cohort studies and randomized controlled clinical trials. There are a series of formulae to explore minimum sample size relating to varying research questions, the study's hypothesis and types of outcomes including odds ratio, risk ratio, and rate ratio; it can easily be explored using statistical packages. The important task is to understand the research question, hypothesis, and outcomes correctly otherwise, even use of packages stands to be deceptive; one must choose the methods of sample size exploration accurately.

Sample Size for Cell-Based Studies and Animal Studies

As such clinical studies aiming to a generation of newer reliable evidence through randomized controlled trials, begin with cell-based, animal study, Phase I trial, Phase II trial, Phase III trial, and Phase IV trials. There are ways[2] to explore sample size for cell-based and animal-based clinical studies. But as such, they do not involve a serious sample size exploration. As a rule of thumb, under cell-based study, I think considering a feasible 10 to 30 samples in each group may be a better proposition giving useful related information. On the other hand, for obvious reasons, the smallest animals need to be sacrificed in an animal study. Accordingly, one must decide 3 to 7 animals in each group to derive suggestive preliminary findings on the topic.

Sample Size for a Clinical Trial

There is a host of literature on this topic.[3,4] There must be a realistic approach towards sample size exploration. Also, the problem of small sample size may not be ignored. In fact, out of four phases of the randomized controlled trial, sample size exploration is involved only in Phase III. As such, Phase I is conducted on volunteers with a major focus on toxicity. As a rule of thumb, 10 to 30 patients might be involved[3]. Phase II trial is conducted again, focusing on toxicity along with suggestive efficacy. Under this, as a rule of thumb, 100 to 200 patients need to be covered under each arm involving selected drugs/ doses through Phase I study[3]. It may be worth mentioning here that postgraduate (e.g., MD/MS/DM/

PhD) students may often not be able to cover explored number of patients. They land up doing a pilot Phase II trial covering a feasible number of patients in each arm, of course in consultation with their preceptors/ guides. Although they do such dissertations from a training point of view, they need to have perfect knowledge about every step under research methodology, including sample size exploration, and then decide what is feasible for them. Further, although randomization is not mandatory under Phase II trial, they need to learn and do randomization even under Phase II trial adopting an approach (e.g., permuted block randomization) ensuring equal number in each arm. This will help in gaining possible optimal power of the study with a considered sample size.

For Phase III trial, as stated earlier, basic principles of sample size exploration will remain same as in case of comparative study/hypothesis testing. For example, is the new drug efficacious? In this case, one group of patients will receive 'new drug' and another group a' placebo'. The required inputs after due thought under this objective will involve normal deviate at considered level of confidence; normal deviate at the considered power of the study; proportion of cured patients in first group and proportion of cured patients in second group; and clinically acceptable effect size. In this case, level of confidence[5,6] may be defined as probability/ chance of correctly concluding that proposed drug is not efficacious. On the other hand, power of the study[5, 6] may be defined as probability/chance of correctly concluding that drug is efficacious. In contrary to the conventional consideration of level of confidence and power of the study as 95 and 80% respectively, under randomized trials, it is advisable to increase them further, like 99 and 90%, respectively. Once we increase them, the calculated sample size for each group will move higher. Further, contrary to the study's power, normal deviate value at a specific level of confidence remains higher in case of two-sided hypothesis of the study (*e.g.,* superiority trials) than that in case of a one-sided hypothesis (*e.g.,* non-inferiority trials). As such, clinically acceptable effect size, in this case, needs to be debated and decided in consultation with public health experts. As obvious, being in denominator, it is inversely correlated with required minimum sample size. In contrast, being in numerator, the higher the pooled variance and the higher the required minimum sample size. As obvious, if outcome is taken as relative risk/ rate ratio instead of proportion of cure rates, sample size exploration formulae might change but basic principle will remain same while exploring sample size.

What if there are Multiple Objectives/ Outcomes?

The sample size must be explored in view of primary objectives/ outcomes. If one primary objective but more than one outcome, sample size needs to be explored for each outcome, and the largest one will serve the purpose for each outcome. If more than one objective, similar approach needs to be adopted. For each objective, exploration needs to be done for each involved outcome(s). Further, as per the explored list, largest explored sample size must be considered as a minimum sample size for the study.

What if there is no literature on the topic?

The required minimum sample size is explored considering each of the required inputs (described above), especially risk measures and effect size, more objectively. In the absence of ways to decide inputs more objectively, it is better to begin carrying out a pilot study/Phase II trial and then use the observed results in exploring the sample size for Phase III trial. It may be worthwhile to mention here that in case of larger explored minimum sample size, if necessary, multi-centric study may be planned to cover the sample size smoothly. However, in this case, one needs to use available guidelines to ensure quality data leading to reliable evidence.

What if Results are for away from Considered Inputs?

It is better to explore power analysis considering the covered sample size and the observed results as inputs, if it emerges to be higher or similar to the considered one, no problem. Otherwise, in view of the requirement of retaining optimal power, the study will fall in the category of a pilot study. It will accordingly restrict the derived summary of the results and its implications.

Sample Size vs. Conclusive/Pilot Study

A study completed with all scientific considerations, including the required minimum sample size, may be conclusive. In other words, it may allow concluding the findings and its related implications. Otherwise, studies remain pilot study which may not allow making conclusions. The results will remain suggestive/ indicative.

Minimum Sample Size vs. Larger Sample Size vs. Small Sample Size

Often, the clinical colleagues go with a myth that consideration of larger sample size may make the smaller effect size as significant. But, there is hardly any study covering larger sample size than the required minimum sample size. Even if there is study covering larger sample size, there will be no much harm. The results will be more precise and accurate. In other terms, study covering larger sample size may not be deceptive. For example, while comparing new drug with standard drug, if standard drug emerges to be superior, patients will continue with same drug as earlier. Even if new drug emerges to be effective, it will be based on larger sample size with higher precision. No scope of much harm to the patients. On the contrary, studies covering smaller than minimum required sample size will have lower power and provide no significant results. In this case, although study is not planned for, there is chance of misinterpretation that new drug is as effective as standard drug. As a result, patients will be likely to be treated with inferior drugs and suffer further.

Summary

For a conclusive study, if required, one must explore the required minimum sample size to answer the proposed research question appropriately. If not feasible to cover explored sample size due to time limit/resources, same topic can be allotted to successive students under the guidance of same guides/consultants to ensure continuing same methodologies. Under such practice, if desired, there might be publication by every student using data of his/her study period, of course in a suggestive manner only. Once required minimum sample size is cumulatively completed by few students, data can be pooled together and analyzed, with due authorship to everyone. Otherwise, in case of stopping as a pilot study, no conclusion needs to be drawn. Also, in this case, research article needs to be written in a suggestive/ indicative manner leaving scope for future study. The same principle applies in case of multi-centric studies. It may be worthwhile to mention here that regardless of conclusive/pilot study, one needs to write how a study was done, not otherwise.

There is need to adopt a multidisciplinary approach to ensure quality research. To emphasize the importance of sample size exploration for clinical studies, it will not be an exaggeration to state that if a clinical diagnosis by a clinician goes wrong, only one patient is likely to suffer. On the contrary, if statistical diagnosis goes wrong regarding drug efficacy, a series of patients will continue suffering. Further, morality comes before research methodology. Further, there is no shortcut method to establish reliable evidence which can guide clinical/ public health practice. In summary, in consultation with a biostatistician[7], consideration of the required minimum sample size to accurately answer a specific research question in medical research is inescapable.

Acknowledgment

The author expresses his thanks and gratitude to each and every editorial board member of this esteemed journal for expressing the need to write the CME Series on the role of biostatistics in medicine.

References

1. Dwivedi SN, Singh R. On assessing the child spacing effect of breastfeeding using Cox- Proportional hazards models. Demography India. 2003; 32(2): 160-170.
2. Lazic SE, Clarke-Williams CJ, Munafo MR. What exactly is 'N' in cell culture and animal experiments? PLoSBiol.2018April4; 16(4):e2005282.doi:10.1371/journal.pbio.2005282. PMID: 29617358;PMCID: PMC5902037.
3. Pocock Stuart J. Clinical Trials: A Practical Approach. John Wiley & Sons Ltd. 2013 July. Online ISBN: 9781118793916.
4. Sundaram KR, Dwivedi SN, Sreenivas V. Medical Statistics: Principles and Methods. Wolters and Kluwer (Health). New Delhi. 2015 (Second Edition). ISBN: 978-9351293217.
5. Armitage P, Berry G, Matthews JNS. Statistics in Medical Research. Fourth Edition. Wiley Inter Science. 2008. ISBN: 9780632052578.
6. Altman Douglas G. Practical Statistics for Medical Research. Chapman & Hall/ CRC. 1991. ISBN 0-412- 27630-5.
7. Dwivedi SN. Need of Strong Biostatistics Component in Medical Colleges in India: Time to Accept the Overdue Reality. Editorial, Central India Journal of Medical Research. 2022; 1(2):2-4.