TEXNİKA ELMLƏRİ

TECHNICAL SCIENCES

DOI: https://doi.org/10.36719/2663-4619/102/198-213

Habib Elsayir
Umm Al- Qura University
habibsayiroi@Yahoo.com
Ibrahim Hassan Alkhairy
Umm Al-Qura University
ihkhairy@uqu.edu.sa

ESTIMATING THE SAMPLE SIZE FOR EPIDEMIC AND MEDICAL RESEARCH

Abstract

This study aims at developing a method of estimating an optimal sample size for use in health and medical research and studies. The paper is designed as a tool that a researcher could use in planning and conducting good quality research and gives a discussion of various aspects of sample size consideration in medical research. The paper also covers the essentials in calculating power and sample size for a variety of applied study designs. Sample size computation for survey type of studies, observation studies and experimental studies based on means and proportions or rates, sensitivity – specificity tests for assessing the categorical outcome are presented in detail. Recently, considerable interest has been focused on medical research after the beginning of COVID-19 pandemic. The resulting literature is scattered over many sources. This paper aimed at giving some contributions in this field.

Keywords: *clinical significance, confidence interval, research design, sample size estimation, statistical power, type 1 and type 2 error*

Həbib Elsayir
Umm Əl-Qura Universiteti
habibsayiroi@Yahoo.com
İbrahim Həsən Əlxairi
Umm Əl-Qura Universiteti
ihkhairy@uqu.edu.sa

Epidemik və tibbi tədqiqat üçün nümunə ölçüsünün qiymətləndirilməsi

Xülasə

Bu iş sağlamlıq və tibbi tədqiqat və tədqiqatlarda istifadə üçün optimal nümunə ölçüsünü qiymətləndirmək üçün bir metod hazırlamaq məqsədi daşıyır. Məqalə tədqiqatçının keyfiyyətli tədqiqatın planlaşdırılması və aparılmasında istifadə edə biləcəyi bir vasitə kimi nəzərdə tutulmuşdur və tibbi tədqiqatlarda nümunə ölçüsünün nəzərə alınmasının müxtəlif aspektlərini müzakirə edir. Sənəd həmçinin müxtəlif tətbiqi tədqiqat dizaynları üçün gücün və nümunə ölçüsünün hesablanmasında əsasları əhatə edir. Tədqiqat tipli tədqiqatlar üçün nümunə ölçüsünün hesablanması, vasitələr və nisbətlər və ya dərəcələrə əsaslanan müşahidə tədqiqatları və eksperimental tədqiqatlar, kateqoriyalı nəticənin qiymətləndirilməsi üçün həssaslıq - spesifiklik testləri ətraflı şəkildə təqdim olunur. Bu yaxınlarda, COVID-19 pandemiyasının başlamasından sonra tibbi tədqiqatlara böyük maraq artır. Nəticədə ortaya çıxan ədəbiyyat bir çox mənbələrə səpələnmişdir. Bu yazı bu sahədə bəzi töhfələr vermək məqsədi daşıyırdı.

Açar sözlər: *klinik əhəmiyyət, etibar intervalı, tədqiqat dizaynı, nümunə ölçüsünün qiymətləndirilməsi, statistik güc, tip 1 və tip 2 xətası*

Introduction

The sample size is the number of observations or specimens required in a study. A too-large sample is merely a waste of resources and time and on the other hand, too small a sample fails to produce conclusive and reliable results, N.Gopi Chander (2017). Therefore, depending on the study design and the outcome, which is estimated before the start of the study, a researcher needs to estimate the optimum sample size by the scientific method to produce reliable results, which can serve as a strong foundation for evidence-based practices. The despotic or inadequate calculation of sample size can affect the research design and its significance. The larger size can lead to ethical concerns, time-wasting, and financial costs, and a smaller sample size influences the power of the study. Very recently, it was observed that many medical-related researchers were rushing to conduct their research on COVID-19 and related diseases research and they have to meet the challenge of necessary statistical considerations regarding sample size calculation to provide essential information concerning sample size determination, including the level of significance, the desired power, and the estimated effect size to achieve the desired research power (see for instance Zhang Q (Zhang, Wang, Qi, Shen, Li, 2019); Zhou YH (Zhou, Qin, Lu, 2020), Yelin I (Yelin, Aharony, Shaer-Tamar, Argoetti, Messer, Berenbaum, 2020). Also, much literature about the topic can be found in Russel (Russel, 2001), Bellera et al (Bellera, 2012) Manski (Manski, 2016) and Johnston (Johnston, Lakzadeh, Donato, 2019).

It has been observed that many published research lacks clinical relevance and lacks clarity about the sample size estimation as well as less power due to inappropriate sample size, see (Paul, 2020; Gautret, Lagier, Parola, 2020; Hogan, Sahoo, Pinsky, 2020; Wu, Xia, Li, 2020). Sample size estimation is also essential to know the feasibility of study in terms of required cost and time. Paul H. Lee (Paul, 2020) reviewed some research trials about COVID-19 which were published between Jan. 1, 2020, and Mar. 25, 2020, and indexed in PubMed, and assessed the quality of their sample size calculation. He identified a total of 374 articles and reached to conclude that, in general, the quality of sample size calculation was not acceptable. Some of these studies did not justify the sample size, others have problems and mistakes concerning Cohen's *d* effect size and the state of the null hypothesis assumption of the control group. For instance, in Gautret et al. (Gautret, Lagier, Parola, 2020) the authors only reported the effect of the treatment (hydroxychloroquine) group ("Assuming a 50% efficacy of hydroxychloroquine in reducing the viral load at day 7") but the effect of the control group was missing. However, in the time of rapid disease outbreak such as COVID-19, researchers are working around the clock to examine the effectiveness of potential treatments, and inappropriate sample size calculation will lead to adverse consequences. The results showed that the method for attaining the desired precision of expected width provides satisfactory results only when sample sizes are large. Arif Habib et al (Arif, 2014), Manski (Manski, Tetenov, 2016), Manski (Manski, Tetenov, 2019) articles addressed the choice of smallest effect, sample size with various designs, the effect of validity and reliability of dependent and predictor variables, sample size for comparison of subgroups and individual differences and responses, sample size when adjusting for subgroups of unequal size. Some practical guidelines for effective sample size determination were found in Lenth (Lenth, 2017). Hopkins (2018) presented a spreadsheet using two new methods for estimating sample size for a study designed to make an inference about real-world significance, which requires interpretation of the magnitude of an outcome, based on acceptable uncertainty defined either by the width of the confidence interval or by error rates for a clinical or practical decision arising from the study. Aimed to develop and present accurate and strict approaches for sample size computation using a real-world case study, Karissa M.et al (2019) presented formal guidance approach on sample size calculations for retrospective burden of illness studies, which was designed for practical application in real-world review studies. Sharma et al (2020) article covered different formulas of sample size calculation for different types of variables measured in distinct study designs, namely descriptive, epidemiological, comparative, and interventional research studies. Aimed to develop and present accurate and strict approaches for sample size computation using a real-world case study, Karissa M.et al(2019) presented a formal

guidance approach on sample size calculations for the retrospective burden of illness studies, which was designed for practical application in real-world review studies. Karissa M. et al(2019) presented sample size formulae for parameters that are of frequent interest in the context of a burden of illness study, see also Johnston, K.M(2019). Sharma et al(2020) article covered different formulas of sample size calculation for different types of variables measured in distinct study designs, namely descriptive, epidemiological, comparative, and interventional research studies. Althubaiti (2023) also, introduced a simple practical guide for health researchers aim to help researchers and health practitioners interested in quantitative research.

The rest of this paper is organized as follows: Chapter 2 is devoted to sample size approaches, formulas and calculation procedures, and the relation with confidence level and power analysis. Numerical illustrations and examples have been done in chapter 3, followed by results discussion in chapter 4, while chapter 5 is devoted to conclusions that have been reached in this research.

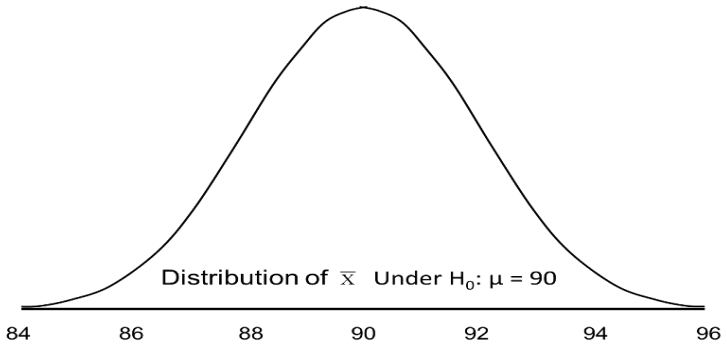
2.Methodology of Sample Size Estimation.

2.1.Sample Size Approaches.

A popular approach to determining sample size involves studying the power of a test of hypothesis including specifying a hypothesis and significance test on a parameter, specifying the effect size for the scientific interest, obtaining historical parameter estimates to be used to compute the power function of the test and specifying a target power value as an objective value of the test, see Russel (2001).

Several equations are used to determine the minimum number of subjects that need to be included in a study to have sufficient statistical power to detect an effect in medical research. statistical power is determined by various variables, such as the variance, and treatment effect size ,see Morris (2012). For instance, the clinically considered **i**mportant difference is determined by clinical experience. The smaller the clinically important difference, the more difficult it will be to prove statistically, and the larger the sample size necessary.

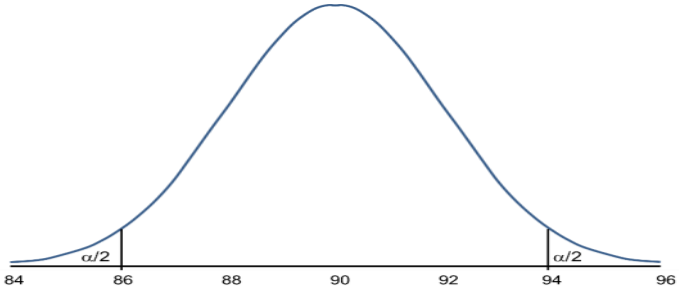
In each test of hypothesis, two errors can be committed, a Type I error which refers to the situation where we incorrectly reject H_0 when in fact it is true, whereas the second type is called a Type II error and is defined as the probability we do not reject H_0 when it is false. In hypothesis testing, we usually focus on power, which is the probability that a test correctly rejects a false null hypothesis. A good test is one with a low probability of committing a Type I error (i.e., small α) and high power (i.e., small β, high power).Suppose we want to test the following hypotheses at α=0.05: H_0: $\mu = 90$ versus H_1: $\mu \neq 90$. Suppose a sample of size n=100 is selected, and the standard deviation of the outcome is σ =20. Then a test statistic is computed and compared to an appropriate critical value. If the null hypothesis is true (μ=90), then we are probably to get a sample with mean close in value to 90 and it is likely to observe any sample mean as shown in figure (1) under H_0.



Figure(1)Mean distribution under null hypothesis

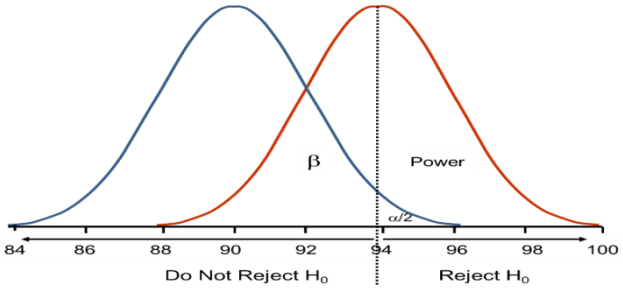
To construct the decision rule for our test of hypothesis, we choose critical values based on α=0.05 and a two-tailed test. So, the decision rule is as follows: Rejection area for test H_0: $\mu = 90$

versus H_1: $\mu \neq 90$ at α =0.05. The areas in the two tails of the curve in figure(2)represent the likelihood of a Type I Error, α= 0.05.



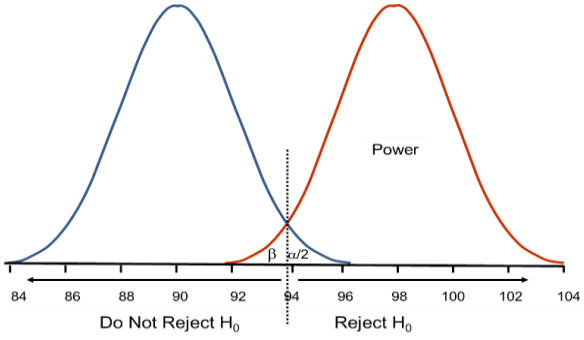
Figure(2)Area for type I error

Now, if the alternative hypothesis, H_1, is true (i.e., $\mu \neq 90$) ,i.e. ,the mean = 94. The figure (3) shows the distributions of the sample mean under H_0. The sample mean values are shown along the horizontal axis figure (3).



Figure(3) Distribution of $\bar{x}$ Under H_0

If the actual mean is 94, then the $H_1$is true. For the test, α is set at 0.05 and reject H_0 if the observed sample mean exceeds 93.92 (see the upper tail of the rejection area). The critical value (93.92) is exhibited by the vertical line. The probability of a Type II error is denoted β, and β = P(Do not Reject H_0 | H_0 is false), β is shown in the figure (3) (where we do not reject H_0). Power, (1- β is shown in the figure as the area under the rightmost curve (H_1) to the right of the vertical line (where we reject H_0). Note that β and power are related to α, the variability of the outcome and the effect size. From the figure above we can see what happens to β and power if we increase α.The figure (4) shows the same structure for the case where the mean under the alternative hypothesis is 98,where distribution of $\bar{x}$ under H_0: $\mu = 90$ and under H_1: $\mu = 98$.



Figure(4) Distribution of $\bar{x}$ Under H_0 and H_1

It is observed that there is much greater power when there is a greater difference between the mean under H_0 as compared to H_1 (i.e., 90 versus 98). A statistical test is much more probable to

reject the H_0in favor of the H_1 if the true mean is 98 than if the true mean is 94. In this situation, it is seen that there is little overlap in the distributions under the H_0 and H_1 hypotheses.

2.2.The formulas for sample size calculations:

The formulas for calculating the required sample size based on the nature of the population data, whether the data collected is to be of a categorical or quantitative nature require knowledge of the variance or proportion and a determination to the maximum desirable error, as well as the acceptable type I error risk (e.g., confidence level). It is possible to construct a table that suggests the optimal sample size, given a population size, a specific margin of error, and a desired confidence interval, see Harza A. (Harza, 2017).

2.2.1.Sample size calculations for categorical outcomes (e.g., treatment patterns):

When considering treatment distributions in a population, with a binomial distribution assumption of (n, p), in which the n stands for the sample size and p represents the probability of receiving the treatment, the following formulas are hold (Karissa M. et al(2019), Arif Habib et al(2014), and Sharma et al. (2020)):

$$\mp 1.96\sqrt{\frac{P(1-P)}{n}} \text{.....(1)}$$

The probability of not receiving the treatment is $(1-p)$, and therefore the probability of all n patients not receiving the treatment is $(1-p)^n$, such that the probability of observing at least one patient receiving the treatment. These formulae can be used to define a sample size that ensures all key treatments will be observed, and that the proportions can be estimated within desired precision. To utilize them to generate sample size requirements, limited a priori data are required; n can be selected to yield acceptable values for both quantities. Because the required sample size n increases as p moves further from 0.50, p can be defined to be the most extreme proportion that would be of interest (e.g., a rare event of 1% portion of the population).

2.2.2.Sample size calculations for continuous outcomes (e.g., costs):

When regarding medical costs, assuming that the mean cost μ is normally distributed and the population standard deviation is σ, the precision associated with a particular sample size can be recognized by the width W of the 95% confidence interval (CI): If an estimate of σ is available, e.g. based on published evidence for another indication, then the width of the CI can be expressed for the maximum feasible sample size n. Alternatively, Eq. (1) can be rearranged so that the required n can be calculated as:

$$n = \frac{1.96^2\sigma}{w^2} \text{.....(2)}$$

To receive 95% CIs for a desired value ± W:

$$\mp 1.96\frac{\sigma}{\sqrt{n}} \text{.....(3)}$$

Occasionally, estimates of the population variance σ^2 aren't available, making sample size calculations challenging. Since σ is unknown, an option available is to consider the coefficient of variation c_v. Based on this, σ can be expressed as $c_v \times \mu$, and for assumed values of c_v and μ, n can be estimated for σ. Then the width of a 95% CI can be expressed as:

$$\mp 1.96\frac{CV\mu}{\sqrt{n}} \text{.....(4)}$$

For cost estimation, based on Eqs. (2) and (3), assuming that an estimate for the standard deviation is not available, an estimate of c_v can instead be used. In practice, a range of possible

values can be considered, based on cost and any a priori knowledge regarding the population for health resource utilization, expected range of disease severity, and care vs. high-cost acute treatment such as inpatient stays.

2.2.3.Estimation of Sample Size for Cross-sectional or Descriptive Research Studies

These studies or surveys are generally conducted to find out, observe, describe, and document aspects of a situation as it naturally occurs. It is not used to identify the causation of something, such as a reason for an epidemic. A researcher might collect cross-sectional data on past alcohol habits and current diagnoses of liver disease, for example.

Following Richard, A (Richard, Zeller, 2007), Bellera et al(2012), and Sharma SK(2020) the sample size calculations were given below:

Sample size in case data is on nominal/ordinal scale and proportion is one of the parameters:

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}}\right)^2 * (P)(q)}{d^2} \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

n = Desired sample size

Z1−α/2 = Critical standard value for the corresponding level of confidence.

(At 95% CI or 5% level of significance (type-I error) it is 1.96 and at 99% CI it is 2.58)

P = Expected prevalence or based on previous research

q = 1-p ,and d = Margin of error or precision.

2.2.4.Sample Estimation for Case–control Studies.

It is a study that determines the cause and effect to see whether exposure is correlated with an outcome or not. By way of explanation, it determines wherever an exposure is correlated with an outcome (i.e., disease or condition of interest). It is a type of observational study in which commonly assumed causation is studied among two groups differing in outcome. For example, a case–control research to find out the relationship between alcohol and liver disease.

Example I: Sample size, when proportion is parameter of the study or data are on nominal/ordinal scale:

$$n = \frac{\frac{(r+1)}{r} * (P(1-P)\left(Z_{1-\beta} + Z_{1-\frac{\alpha}{2}}\right)^2}{(P_1 - P_2)^2} \ldots\ldots\ldots (6)$$

Where:

r = Control to cases ratio (1 if same numbers of subject in both groups)

p = Proportion of population = (P1+P2)/2

Z1-β is the desired power (0.84 for 80% power and 1.28 for 90% power)

z1-α/2 = Critical value for the corresponding level of confidence.

(At 95% CI or 5% type I error it is 1.96 and at 99% CI or 1% type I error it is 2.58)

P1 = Proportion in cases and P2 = Proportion in controls.

Sample size n in case data is on interval/ratio (quantitative) scale and mean as a parameter of the study:

$$n = \frac{\frac{(r+1)}{r} * \sigma^2\left(Z_{1-\beta} + Z_{1-\frac{\alpha}{2}}\right)^2}{(d)^2} \ldots\ldots\ldots\ldots (7)$$

Where σ = SD which is based on a previous study or pilot study d = Effect size (difference in the means from previous studies or pilot study)

2.2.5.Sample size estimation for comparative studies.

It is the study design in which comparison is done between two or more groups based on selected attributes such as knowledge, perception, and attitude. A multidisciplinary approach is best

used for this type of research. Sample size in case data is on nominal/ordinal scale and proportion is the parameter of the study:

$$n = \frac{(P_1(1-P_1)+P_2(1-P_2)}{(P_1-P_2)^2} * C \ldots\ldots\ldots(8)$$

n = Sample size for one group that we need to find out

p1 and p2 = Proportion of two groups

C = Standard value for the corresponding level of α and β selected for the study. It is as follows: Sample size in case data is on interval/ratio scale and mean is parameter of the study.

$$n = \frac{(\sigma^2{}_1+\sigma^2{}_2)^2 * \left(Z_{1-\beta}+Z_{1-\frac{\alpha}{2}}\right)^2}{(d)^2} \ldots\ldots\ldots..(9)$$

d = difference in means of two group (effect size)

σ1 = SD of Group 1,σ2 = SD of Group 2

Z1-β = It is the desired power ,and z1-α/2 = Critical value and a standard value for the corresponding level of confidence. (At 95% CI it is 1.96 and at 99% CI, or 1% type I error it is 2.58).

2.2.6.Sample Size Estimation for Experimental Studies

Experimental studies or randomized controlled trials are the studies in which researcher artificially manipulates variables under the study. Randomization and control group are important aspect in these types of studies. In this investigator provides intervention and study its effect and compare in experiential and control group.

Example I: Sample size to rule out the difference (effect size) among two groups (based on difference in proportion or for dichotomous nominal/ordinal variables)

n = Sample size for each group

d = Difference in means of two treatment effect

Zx = Standard value for a one or two-tailed

σ2 = SD of Group 2

$\delta 0$= Acceptable margin of error

S2 = Pooled SD (both comparison groups)

p = Response rate of standard intervention

p0 = Response rate of new intervention.

a. Non-inferiority trial:

$$n = \quad 2 * \{\frac{Z_{1-\beta}+Z^2{}_{1-\frac{\alpha}{2}}*P}{\delta_0}\} * (1-P); \ldots\ldots\ldots..(10)$$

or

$$n = \quad 2 * \{\frac{Z_{1-\beta}+Z^2{}_{1-\frac{\alpha}{2}}}{\delta_0}\} * SD^2 \ldots\ldots\ldots(11)$$

b. Equivalence trial

$$n = \quad 2 * \{\frac{Z_{1-\beta}+Z_{1-\frac{\alpha}{2}}}{\delta_0}\} * P(1-P) \; ; \ldots\ldots..(12)$$

or

$$n = \quad 2 * \{\frac{Z_{1-\beta}+Z^2{}_{1-\frac{\alpha}{2}}}{\delta_0}\} * SD^2 \ldots\ldots\ldots(13)$$

c. Clinical superiority trial

Sample size n

$$n = \quad 2 * \left\{\frac{Z_{1-\beta}+Z^2{}_{1-\alpha}}{\delta-\delta_0}\right\} * P * (1-P) \; ; \ldots\ldots..(14)$$

or

$$n = \quad 2 * \left\{\frac{Z_{1-\beta} + Z^2{}_{1-\alpha}}{\delta - \delta_0}\right\} * SD^2 \ldots\ldots\ldots (15)$$

Sample size to rule out the difference (effect size) among two groups (based on difference in the mean or for continuous variables).

Estimates of the true prevalence of Covid-19 can be made by random sampling in the wider population. Ola Brynildsrud(2020) used simulations to explore confidence intervals of prevalence estimates under different sampling intensities and degrees of sample pooling and based on simulation of the effect sample pooling on prevalence estimates under different settings for true prevalence. Starting by generating a population of *p* individuals and then let everyone have *p* probability of being infected at sampling time. If *n* is the number of patient samples collected from the population, and the number of patient samples that are pooled into a single well is denoted by k, then the total number of pools are thus $\binom{n}{k}$.

Pooled sampling can be used to efficiently assert freedom from disease with a certain probability. If the population is free from the disease, then we find no true positive specimen in our sampling. Ola Brynildsrud (Brynildsrud, 2020) calculated how many samples we need to take from a population with prevalence to ensure that the probability of sampling *p* at least one single positive patient is α or higher. The needed number using the formula of Christensen and Gardner (Christensen, Gardner, 2000) is:

$$n \geq \frac{\log\big(1 - (1-\propto)\big)}{\log(\theta(1 - \mathrm{p}) + (1 - \eta)\mathrm{p})} \ldots\ldots\ldots (16)$$

Where n:number of patients samples from population,θ:specified test value and η is sensitivity value (for example 0.95).

For tests with perfect specificity, we do not have to worry about false positives, and if any pools come out as positive, we classify the population as not free from disease. The formula of Christensen and Gardner (2000) can be expanded to the case with pooled sampling:

$$m \geq \frac{\log\big(1 - (1-\propto)\big)}{\log(1-p)^k\theta + (1 - (1-p)^k(1 - \eta))} \ldots\ldots\ldots (17)$$

Where:

$m = \binom{n}{k}$: total number of pools.

3.Numerical Analysis:

There are different equations that can be used to calculate sample size and confidence intervals depending on factors such as whether the standard deviation is known or smaller samples (n<30) are involved, among others. Most commonly however, the population is used to refer to a group of people, whether they are the number of patients in a hospital, the number of patients within a certain age group of some geographic area, or the number of patients in a hospital at any given time. Based on different formulas, we present the recommended sample size and its relationships between related factors (Corman, Landt, Kaiser, Molenkamp, Meijer, Chu, 2020).

The Sample Size Calculator uses the following formulas:

1. $n = z^2 * p * (1 - p) / e^2$

2. n (with finite population correction) = $[z^2 * p * (1 - p) / e^2] / [1 + (z^2 * p * (1 - p) / (e^2 * N))]$

Where:

n is the sample size; **N** is the population size.

z is the z-score associated with a level of confidence,**p** is the sample proportion, expressed as a decimal,and **e** is the margin of error, expressed as a decimal,

Numerical Example: Let's say we want to calculate the proportion of COVIT19 patients who have been discharged from a given hospital who are happy with the level of care they received while hospitalized at a 90% confidence level of the proportion within 7%. Then we ask :What sample size would we require (Eur, 2020)?

The sample size (n) can be computed using the following formula:

$n = z^2 * p * (1 - p) / e^2$

where Z= 1.645 for a confidence level (α) of 90%, p = proportion (expressed as a decimal), E = margin of error.

Z = 1.645, p = 0.5, E = 0.07

$n = 1.645^2 * 0.5 * (1 - 0.5) / 0.07^2$

n = 1.3530125 / 0.0049 = 276.125 ≈ 276 patients.

Margin of Error vs. Sample Size:

Margin of error indicates the extent to which the outputs of the sample population are reflective of the overall population. The lower the margin of error, the nearer the researcher is to have an accurate.

Example:

Given: Confidence level (**α**),Margin of Error **(E)**: %, Population Proportion **(p)**: %. Population Size **(N)** (optional).The sample size (n) is calculated according to the formula:

$n = [z^2 * p * (1 - p) / e^2] / [1 + (z^2 * p * (1 - p) / (e^2 * N))]$ Where: z = 1.96 for a confidence level (α) of 95%, p = proportion (expressed as a decimal), N = population size, E= margin of error.

z = 1.96, p = 0.4, N = 1000, E = 0.05

$n = [1.96^2 * 0.4 * (1 - 0.4) / 0.05^2] / [1 + (1.96^2 * 0.4 * (1 - 0.4) / (0.05^2 * 1000))]$

n = 368.7936 / 1.3687936= 269.429,n ≈ 269

The sample size (with finite population correction) is equal to 269.

Table (1) presents the results of some calculations for sample size for a given confidence interval and margin of error, which may be used to determine the appropriate sample size for almost any study. It is noticed that the sample size is larger for a lower margin of error or a higher level of confidence. The margin of error depends on the size and variability of the sample. Naturally, the error will be smaller if the sample size (n) is large, or the variability of the data (Standard deviation) is less. The sample sizes in the following tables presume that the attributes being measured are distributed normally or nearly so. If this assumption is violated, then the whole population must be surveyed (Fritz, Morris, Richler, 2012).

Table (1):Calculate sample size for given CI and margin of error.

Population Size(N)	Confidence Interval % (95)	Margin of Error	Sample Size(n)
100	**95**	**1**	**99**
100	**95**	**0.5**	**100**
1000	**95**	**0.5**	**975**
1000	**95**	**1**	**906**
10000	**95**	**1**	**4900**
10000	**95**	**0.5**	**7935**
10000	**95**	**0.1**	**9897**
1000	**95**	**0.1**	**999**

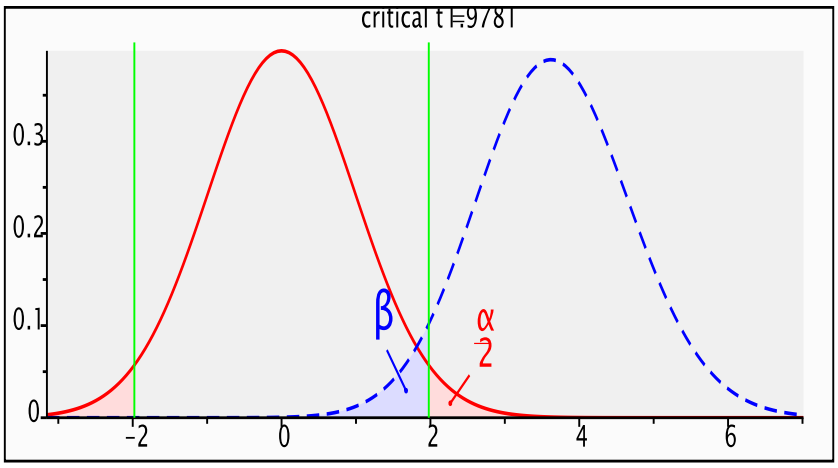
The level of confidence of a sample is expressed as a percentage and describes the extent to which you can be sure it is representative of the target population. For example, a 95% confidence interval does not mean that 95% of the sample data lie within that interval. A confidence interval is not a range of plausible values for the sample, rather it is an interval estimate of plausible values for the population parameter. The proper understanding of CI is not that simple. The true value of the population parameter is fixed, while the width of the 95% CI based on a random sample will also

vary randomly. If repeated random samples of equal size are selected from the population, we will get a corresponding number of 95% CI values, only 95% of them can be expected to combine the population parameter value. Table (2) shows the required sample size for detecting differences from 0.10% to 50%, with 90% confidence and 80% power and conversion rates around 5%. (10% increase for compensation) (Hoekstra, Morey, Rouder, 2014).

Table (2): Calculate sample size for detecting difference.

Required Sample Size				
Difference	Each Group	Total	A	B
0.10%	652,196	1,304,391	5%	5.1%
0.50%	27064	54,129	5%	5.5%
1.00%	7071	14,142	5%	6.0%
5.00%	378	757	5%	10.0%
10.00%	123	246	5%	15.0%
20.00%	44	88	5%	25.0%
30.00%	25	51	5%	35.0%
40.00%	17	33	5%	45.0%
50.00%	12	24	5%	55.0%

In medical research, generally, the researcher should find the optimal sample size by plotting power as a function of effect size and sample size to avoid wasting resources. Figure (6) exhibits the total sample size and power of the test for difference between two dependent means with effect size d=0.3. The effect size represents the lowest difference that would be of significance. It could be the difference in cure rates, or a standardized mean difference or a correlation coefficient. As effect size increases, the type II error decreases. For a specific power, 'small effects' require greater sample size than 'large effects'. Figure (6) shows that an increase in sample size yields greater power (Kane, 2019). The sample size has an indirect effect on power because it affects the measure of variance used to calculate a test statistic (t-test). Since the power of the test to be calculated involves comparison of sample means, one would be more interested in the standard error (the average difference in sample values)than standard deviation or variance. When n, sample size, is large a lower standard error would have been achieved than when n is small. However, when N, the population size is large a smaller beta region would have been achieved than when n is small. The relationship between α and β using t value is shown in figure (5) using t distribution. The figure shows the change in β and power if α is increased. In general, when the α level, the effect size, or the sample size increases, the power level increases (Mizumoto, Kagaya, Zarebski, Chowell, 2020).



Figure(5): The relationship between α and β using t value.

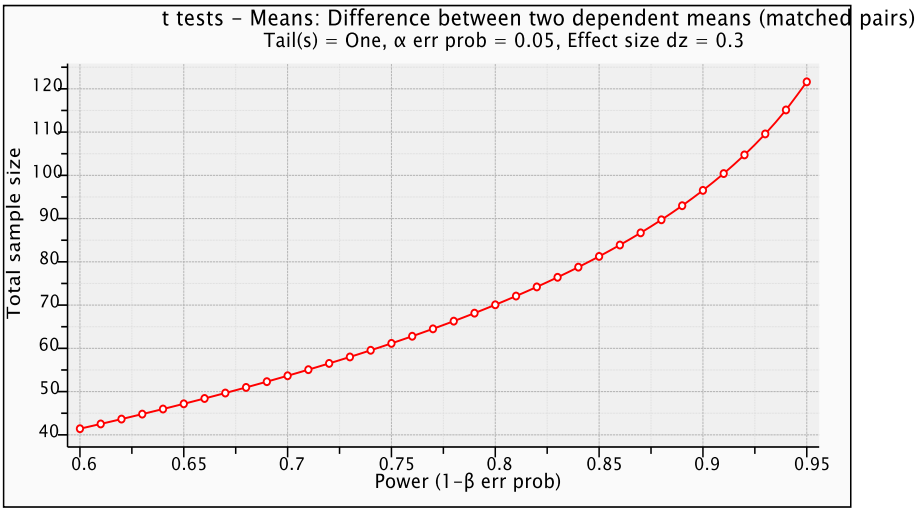


Figure (6): Total sample size and power (Two tailed t test)

Figure (7) shows the two tails t tests correlation plot of α and error probability (point biserial model), where the total sample size values are 10,20 and 30, power at 0.9997392. Considering the alternative hypothesis (H_1), choose a region of rejection such that the probability of observing a sample value in that region is less than or equal to α when accepting H_0. If the obtained sample statistic value falls within the rejection region, the decision is made to reject the H_0. If α is set at 5%, this can be interpreted that in 5%, or one in twenty, the data indicate that "something" exists, while in fact, it does not (Suresh, 2020).

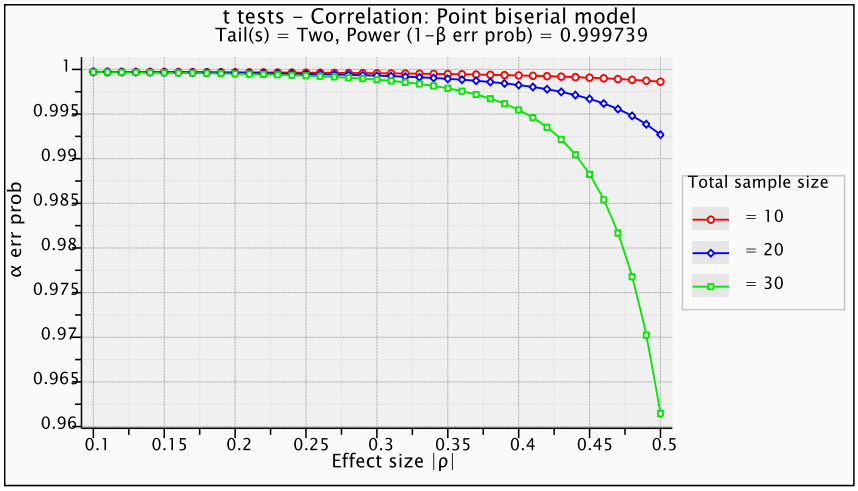


Figure (7): Plot of total sample size, effect size and α prob.

As observed in figure (8),the t test correlation was used (point biserial model) for post-hoc power analysis, given α, sample size and effect size, where effect size (Es(ρ)= 0.3),α=0.05 and the total sample size equals n =300,then power =0.9997392.

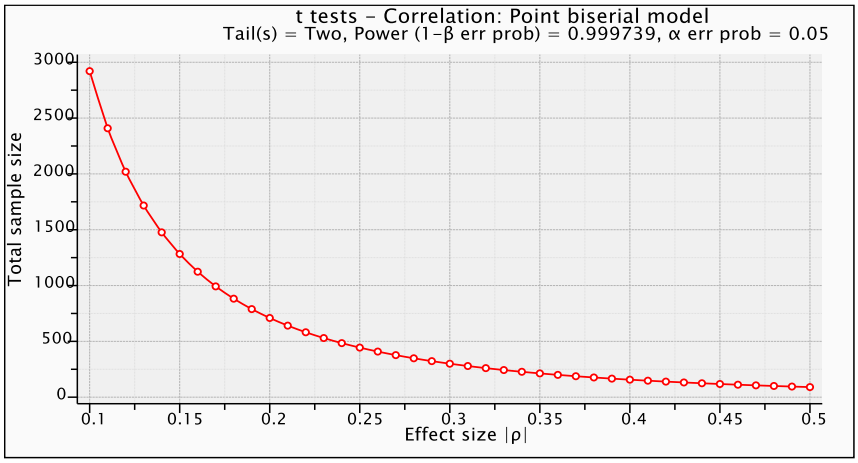


Figure (8):plot of total sample size versus effect size.

The figure(9) illustrates that as effect size(ρ) increases, the error prob(α) decreases.

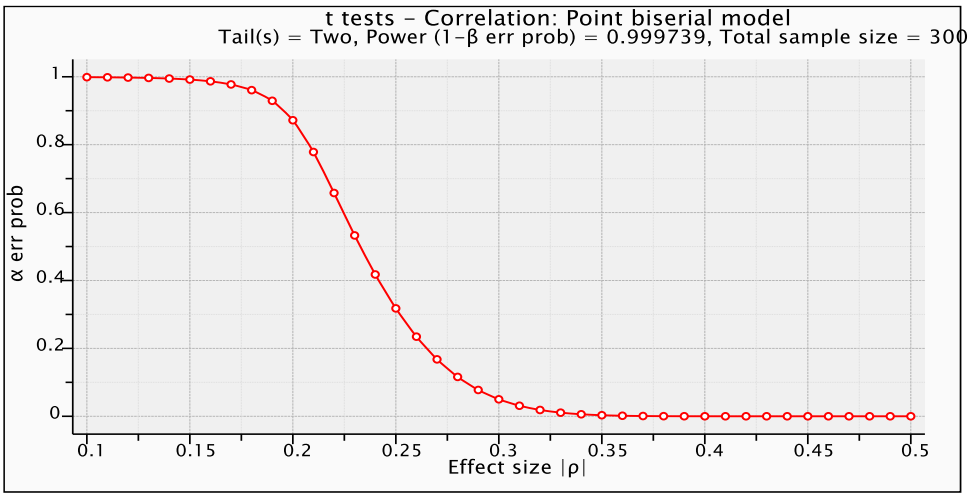


Figure (9):plot of error prob.(α) and effect size(ρ).

When planning a clinical study, the resulting sample size might be too large while the possible resources to conduct such a large study is limited, or ethical reasons may prevent enrolling this many subjects. Reducing the sample size usually involves some compromise, such as accepting a small loss in power.

Evaluating a statistical power of existing medical study:

Let us consider two study groups each received different treatments, Continuous (means). If the primary endpoint was binomial-only two possible outcomes) If the primary endpoint was binomial-only two possible outcomes). E.g., mortality (dead/not dead), pregnant (pregnant/not). Consider the following inputs and after calculating, the results as shown below:

Statistical Parameters

Study Incidence	
Group 1	%
Group 2	%
Number of Subjects	
Group 1	subjects
Group 2	subjects
Type I/II Error Rate	
Alpha	

Results: Dichotomous Endpoint, Two Independent Sample Study

Post-hoc Power	
100% power	
Study Parameters	
Incidence, group 1	30%
Incidence, group 2	70%
Subjects, group 1	500
Subjects, group 2	1200
Alpha	0.05

p_1, p_2 = proportion (incidence) of groups #1 and #2
$\Delta = |p_2-p_1|$ = absolute difference between two proportions
n_1 = sample size for group #1,n_2 = sample size for group #2
α = probability of type I error (usually 0.05)
z = critical Z value for a given α or β
K = sample size ratio for group #2 to group #1
$\Phi()$ = function converting a critical Z value to power

Post-hoc power analysis procedure has been criticized as a means of interpreting negative study results. Because post-hoc analyses are only calculated on negative trials ($p \geq 0.05$), therefore, the analysis will gain a low post-hoc power result, which may lead to misinterpretation as the trial having inappropriate power, hence, instead, 95% CI may be a more appropriate method of calculating statistical power.

Finally, the calculation of the adequate sample size health surveys and studies involves statistical procedures as well as clinical or practical considerations and requires teamwork efforts (also includes biostatisticians) to determine the sample size that will address the research question of interest with adequate precision or power to obtain results that are clinically significant and meaningful.

Discussion.

The formula for sample size calculation alters with the type of study designs. It should be known that all the sample estimates presented represent the largest possible sample size values for the desired level of confidence. Some factors that affect the width of a confidence interval include size of the sample, confidence level, and variability within the sample. As our sample size increases, the confidence in our estimate increases, and then we have greater precision attainable. Higher sample size permits the investigator to increase the significance level of the feedback, since the confidence of the results are potential to increase with higher sample size. It is worthwhile to recall that the confidence interval concept was used to give an answer to an issue in statistical inference results obtained from data that represent randomly selected part of a population. A 95% confidence interval is often used in biological sciences. A much higher level is usually used in the physical sciences, such as engineering field to provide a higher level of precision and eliminate the risks of manufacturing poor-quality products. This can hold for sensitive medical research. Trivial errors in formula, pooling, statistical baseline values, study design, and the outcome measures can lead to erroneous estimation with a great influence on the external validity of the study. In medical research, it is essential sometimes to consider all vital issues, and about 10% of additional samples can be considered to the computed sample size for various consideration. When determining

optimal sample size in medical research, it is important to consider any subgroups of interest, and either target such subgroups directly in sampling strategy or account for expected sample size needs if only the whole population is to be investigated.

Conclusion

An optimal sample size for use in epidemic and medical research studies was introduced in this study. A tool that a researcher could use in planning and conducting good quality research is presented and a discussion of various aspects of sample size consideration in medical research is intensively covered, in addition to the essentials in calculating power and sample size for a variety of applied study designs. Sample size computation for survey type of studies, observation studies and experimental studies based on means and proportions or rates, for assessing the categorical outcome are also presented. Enough details such as the power, significance level, mean or rate for the control group, minimal detectable difference, variance, and dropout rate should be clearly explained in this study. Any other factors that formed the basis of the sample size calculation should also be included. Recently, considerable interest has been focused on medical research after the beginning of COVIT19 pandemic. The resulting literature is scattered over many sources. Therefore, the paper aimed at giving some contributions in this field. Hence, to improve the quality of the sample size calculation of COVID-19 trials and related topics research, it is strongly suggested that all research teams should include a statistician or invite a statistician to evaluate the appropriateness of the sample size calculation. Eventually, the method for estimating sample size in epidemic or any health or medical study should be explained clearly with sufficient detail to permit its use in other protocols later.

References

1. Gopi Chander, N. (2017). Sample size estimation. Jul-Sep; 17(3), pp.217-218. doi: 10.4103/jips.jips_169_17 PMCID: PMC5601499 PMID: 28936033.
2. Zhang, Q., Wang, Y., Qi, C., Shen, L., Li, J. (2019). Clinical trial analysis of 2019-nCoV therapy registered in China. J Med Virol. Accepted. [PMC free article] [PubMed] [Google Scholar].
3. Zhou, Y.H., Qin, Y.Y., Lu, Y.Q. (2020). Effectiveness of glucocorticoid therapy in patients with severe novel coronavirus pneumonia: protocol of a randomized controlled trial. Chin Med J. Article in press. [PMC free article] [PubMed] [Google Scholar].
4. Yelin, I., Aharony, N., Shaer-Tamar, E., Argoetti, A., Messer, E., Berenbaum, D. (2020). Evaluation of COVID-19 RT-qPCR test in multi-sample pools. medRxiv.2020.03.26.20039438.
5. Russel, V.L. (2001). Some Practical Guidelines for Effective Sample Size Determination. Dept of Statistics, University of Iowa.
6. Bellera, C.A. (2012). Calculating Sample Size in Anthropometry in V.R.Preedy (ed.), Handbook of Anthropometry: Physical Measures 3 of Human Form in Health and Disease, DOI 10.1007/978-1-4419-1788-1_1,© Springer Science+Business Media, LLC.
7. Johnston, K.M., Lakzadeh, P., Donato, B.M.K. (2019). Methods of sample size calculation in descriptive retrospective burden of illness studies. BMC Med Res Methodol 19, 9. https://doi.org/10.1186/s12874-018-0657-9.
8. Paul, H.L. (2020). Sample sizes in COVID-19–related research. CMAJ April 27, 192 (17) E461; DOI: https://doi.org/10.1503/cmaj.75308
9. Gautret, P., Lagier, J.C., Parola, P. (2020). Hydroxychloroquine, and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents. Article in press. [PMC free article] [PubMed] [Google Scholar].
10. Hogan, C.A., Sahoo, M.K., Pinsky, B.A. (2020). Sample Pooling as a Strategy to Detect Community Transmission of SARS-CoV-2. JAMA, doi: 10.1001/jama.2020.5445
11. Wu, C.N., Xia, L.Z., Li, K.H. (2020). High-flow nasal-oxygenation-assisted fibreoptic tracheal intubation in critically ill patients with COVID-19 pneumonia: a prospective randomized controlled trial. Br J Anaesth. (article in press) [PMC free article] [PubMed] [Google Scholar].

12. Arif, H. (2014). Design and Determination of The Sample Size in Medical Research. IOSR Journal of Dental and Medical Sciences (IOSR-JDMS) e-ISSN: 2279-0853, p-ISSN: 2279-0861. Volume 13, Issue 5 Ver. VI. (May. 2014), pp.21-31.
www.iosrjournals.org www.iosrjournals.org 21 | Page.
13. Manski, C., Tetenov, A. (2016). "Sufficient Trial Size to Inform Clinical Practice," Proceedings of the National Academy of Sciences, 113(38), pp.10518-10523.
14. Manski, C., Tetenov, A. (2019). "Trial Size for Near-Optimal Treatment: Reconsidering MSLT-II," The American Statistician, 73(S1), pp.305-311.
15. Lenth, R.V. (2017). Some practical guidelines for effective sample size determination. [Last accessed on. Jul 10]. Am Stat, 55, pp.187-193.
Available from: http://www.jstor.org/stable/2685797. [Google Scholar].
16. Harza, A. (2017). Using the confidence interval confidently Thorac Dis; 9(10), pp.4125-4130. doi:10.21037/jtd.2017.09.14.
17. Richard, A., Zeller, Y. (2007). Establishing the Optimum Sample Size. Epidemiology and Medical Statistics. Handbook of Statistics Volume 27, pp.656-678.
18. Christensen, J., Gardner, I.A. (2000). Herd-level interpretation of test results for epidemiologic studies of animal diseases. Preventive Veterinary Medicine, pp.83-106. doi: 10.1016/s0167-5877(00)00118-5.
19. Brynildsrud, O. (2020). Covid-19 prevalence estimation by random sampling in the wider population – Optimal sample pooling under varying assumptions about true prevalence.
20. Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 25.doi: 10.2807/1560-7917.ES.2020.25.3.2000045.
21. Eur, J. (2020). Intern Med. Jul; 77, pp.139-140. Elsevier Public Health Emergency Collection. Published online 2020 Apr 30. doi: 10.1016/j.ejim.2020.04.057 PMCID: PMC7190503 PMID: 32414641.
22. Fritz, C.O., Morris, P.E., Richler, J.J. (2012). Effect size estimates: Current use, calculations, and interpretation. J Exp Psychol Gen. 141, pp.2-18. [PubMed] [Google Scholar].
23. Hoekstra, R., Morey, R.D., Rouder, J.N. (2014). Robust misinterpretation of confidence intervals. Psychon Bull Rev, 21, pp.1157-64. [Crossref] [PubMed].
24. Kane, S.P. (2020). Sample Size Calculator. ClinCalc: https://clincalc.com/stats/samplesize.aspx. Updated July 24. Accessed August 24.
25. Mizumoto, K., Kagaya, K., Zarebski, A., Chowell, G. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship,Yokohama, Japan. Eurosurveillance.
26. Suresh, K.Sh. (2020). How to calculate sample size for observational and experimental nursing research studies?. National Journal of Physiology, Pharmacy and Pharmacology. Vol 10 | Issue 01. DOI: 10.5455/njppp.2020.10.0930717102019.

Received: 15.03.2024 Accepted: 09.05.2024