



# Sample size calculation on web, can we rely on the results?

Alipasha Meysamie<sup>1\*</sup>, Farough Taei<sup>1</sup>, Mohammad-Ali Mohammadi-Vajari<sup>1</sup>, Siamak Yoosefi-Khanghah<sup>1</sup>, Sahra Emamzadeh-Fard<sup>1</sup> and Mehrshad Abbasi<sup>2</sup>

\*Correspondence: [Meysamie@tums.ac.ir](mailto:Meysamie@tums.ac.ir)

<sup>1</sup>Department of Community and Preventive Medicine, Medical Faculty, Tehran University of Medical Sciences, Iran.

<sup>2</sup>Department of Nuclear Medicine, Valiasr Hospital, Tehran University of Medical Sciences, Tehran, Iran.

## Abstract

One of the most frequent concerns of the researchers is sample size calculation. There are many equations to calculate sample size one should be selected correctly based on the type of the study and data type. Any mistake can mislead the researcher and so the study results. Even an overestimation can cause waste of cost, time and material. Either statisticians or other researchers frequently use online advice on sample size calculation. Online sample size calculators have been searched, reviewed and checked for the calculated results based on known formulas for common research objectives. Considering the most common types of health study objectives regarding sample size calculation (estimating proportion or mean, comparison of two proportions or two means), from 60 website or software, only 5 (8.3%) provided all four types of calculations. Overall exact calculation rate was 8.3% and 16.1% per calculation and per site, respectively. Many of the sites just calculate sample size for estimating proportions and most of the results lack the accuracy in calculation.

**Keywords:** Online calculation, sample size, health studies, accuracy

## Introduction

The statisticians, methodologists and non-statistician professional researchers frequently ask this question "how many subjects do I need for my study?" [1]. The sample size affect the feasibility of the research and proper calculation possess significant effect on the power and validity of the results [2]. Currently, the researchers solve most of their problems via online solutions and face to face consultations are limited for the must to do occasions. For internet assisted sample size calculation, the first step is to find a proper site with online calculation capability or downloadable files but the accuracy of calculations is one of the most important matters. There are many online sample size calculation sites with different specifications. Non-statistician users do not like sophisticated processes for calculation and more professional statisticians may need advanced calculations for somehow rare study types. In this review the sample size calculation websites have been listed and tried to be checked how their calculations are the same as basic formulas show.

## Methods

Google search engine have been used with keywords including sample size calculator, sample size calculation, sample size in health studies, sample size web site and sample size software for retrieving related sites. Software informer website (available at <http://software.informer.com>) also used to find downloadable files for sample size calculation. Each site reviewed and data concerning URL, owner, availability of calculation equation(s), providing other types of calculations rather than only sample size calculation e.g. confidence levels or statistical tests collected. Standard formulas for sample size calculation extracted considering

the four most common types of sample size calculation in health studies for estimating proportion of a qualitative variable, a mean of a quantitative variable and comparison of two proportions and two means of two different populations according to detailed review of several statistical resources, such as some biostatistics textbooks [3-8] and articles [9]. Four hypothetical examples considered for calculations (one for each) and the results based on selected formulas calculated (Table 1). Then the results of each site or software compared with the standard results. A difference more than rounding the final results (more than 1) with the standard result, considered as inaccurate (not the same as formula) calculation. The feasibility of site usage also evaluated to determine whether the application of sites was simple and user friendly or professional. To evaluate the complexity, two medical researchers who had some experience in medical research and sample size calculation asked to use sample size calculators to score them as easy, moderate or advanced to use. The sites scored as easy if they required basic input data for calculation and instructed each phrase of calculation clearly for users; the sites which needed detailed information about the study e.g. one tailed or two tailed calculations considered as moderate; and the sites which required very detailed and additional information which were not familiar to routine practices in medical research according to the basis of statistical knowledge as a medical researchers e.g. complex survey analysis or cost analysis considered as advanced. The congruent rating of both observers for sites considered as the ease of use level and if there was any incongruence between their assessment, a third evaluator considered the site review for rating.

**Table 1. Different types of formula for sample size calculation and hypothetical data and results used for assessing the calculation accuracy (z=value on standardized normal distribution for giving probability).**

Type of study	Formula	Hypothetical used data	Calculated result based on formula
Estimation of a proportion	$n = \frac{z_{1-\frac{\alpha}{2}}^2 pq}{d^2}$	p=Given prevalence = 40% α=Type I error = 5% d=Acceptable difference from prevalence=5%	369
Estimation of a mean	$n = \frac{z_{1-\frac{\alpha}{2}}^2 \sigma^2}{d^2}$	σ=Standard Deviation =10 α=Type I error=5% d=Acceptable difference from mean=3	43
Comparison of two proportion	$n = \frac{2 \left( z_{1-\frac{\alpha}{2}} + z_{1-\beta} \right)^2 pq}{(p_1 - p_2)^2}$	α=Type I error=5% β=20% p <sub>1</sub> =Given prevalence for group 1=40% p <sub>2</sub> =Given prevalence for group 2=60%	99 for each group
Comparison of two mean	$n = \frac{\left( z_{1-\frac{\alpha}{2}} + z_{1-\beta} \right)^2 (\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$	α=Type I error=5% β=Type II error=20% σ <sub>1</sub> =Standard Deviation for group 1=7 σ <sub>2</sub> =Standard Deviation for group 2=8 μ <sub>1</sub> =mean for group 1=25 μ <sub>2</sub> =mean for group 2=22	99 for each group

Two sites could not be rated because of unavailability of registration or access for authors [11,52] and some had no calculation regarding to the four common ways considered in evaluation process [12,17,18,27,31,34,52,55,57] (Table 2). STATA software also used for calculations and related results depicted [59].

## Results

The summary of available sample size calculator sites is presented in (Table 2). Among the web-based sample size calculators, some do the statistics online and other provides a downloadable file for calculations [10,11,21,22,24,26,34,53,57,58]. These files vary from a static table in excel [21] or dynamic file for multiple calculations [10]. Most of the sites (25 out of 50) just calculate sample size for estimation of a proportion, and most of them consider a proportion of 50% as a fixed number in formula for calculation and so the results were not accurate for considered example in the study. Only five sites provided all four common types of calculations [16,23,33,45,46]. Considering calculations in comparison with the results of formulas, from 57 different calculations, 18(31.6%) were accurate and based on the sites, only 13(26.0%) sites provides at least one accurate result. From those providing all four types of calculations only one had accurate results for all four calculations based on formulas [23] and the other one [16] provided accurate results for 3 calculations. Some sites provided calculations for specific studies e.g. complex sample surveys [49,57], survival analysis [26,55] and microarray experiments in genetic studies [12] and some provided multi language support [23,37]. From 48 rated sites 28(58.3%) scored as “Easy”, 6(14.6%) as “Moderate” and 10(27.1%) as “Advanced”

according to the complexity of the site usage.

## Discussion

We provided a list of available online sample size calculating sites with brief description and an evaluation of their accuracy in calculations, besides the level of ease of use and complexity.

Online sample size calculators vary from easy and user-friendly to advanced and professional, but what is at the most important matter is “How valid are their calculations?”. According to our findings, there are few sites which provides all common ways of sample size calculation needed for common types of researches and only a few provides accurate calculations. The most important reason for inaccurate results may be due to some fix considerations in many sites e.g. considering a fix proportion of 50% for calculation of sample size in studies when the aim is estimation of prevalence. Some other reasons may be different formulas used in sites rather those we considered in our assessment. It seems providing basic formulas for each calculation, may help researchers to choose more accurate path for sample size calculation. Although in some conditions e.g. comparison of two proportions different formulas will provide almost the same results for sample size calculation. In other conditions like estimation there are somehow more different formulas for sample size calculations and so more apart results. This has been shown in our study that some sites used the formula which considered type II error for sample size calculation for estimation a proportion or a mean [33]. Some sites may use more accurate formulas for calculation but others use inaccurate formulas. For example some sites considers formulas using just one standard deviation for calculation in case of two mean comparison [24,26,46]

**Table 2. Brief description of sample size calculator websites and software, NR=Not Rated, NAP=Not Applicable, NC=No calculation provided.**

Reference	Providing Formulas	Usability	Downloadable File(s)	Type of Calculations Provided	Accurate calculation*	Common Sample Size Calculation Results			
						Estimation		Comparison	
						Qualitative <sup>1</sup>	Quantitative <sup>2</sup>	Qualitative <sup>3</sup>	Quantitative <sup>4</sup>
[10]	No	Advanced	Yes (An excel file)	Calculates one sample mean estimation sample size and also power, error and cost	No	NC	NC	NC	NC
[11]	Yes	NR	Yes	Needs registration for calculation or file download	Not Rated	-	-	-	-
[12]	No	Easy	No	Calculates sample size for microarray experiments in genetic studies	NAP	NC	NC	NC	NC
[13]	Yes	Easy	No	Only for 50% estimate Considers population size Calculates Confidence intervals	No	383	NC	NC	NC
[14]	Yes	Moderate	Yes (Also online)	Only for 50% estimate Considers population size Calculates Confidence intervals	No	383	NC	NC	NC
[15]	Yes	Easy	Yes (Also online)	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[16]	No	Moderate	No	Provides all 4 types of sample size calculations Considers population size One and two tailed calculations	Yes (just 1,2&4)	369	43	107+107	99+99 Uses Variances
[17]	No	Advanced	No	Provides statistical tests for different types of RCT No sample size calculation	NC	NC	NC	NC	NC
[18]	No	Easy	No	Provides Java applets for Different statistical tests and confidence interval calculations No sample size calculation	NAP	NC	NC	NC	NC
[19]	No	Easy	No	Just for a qualitative estimation Considers population size Calculates Confidence intervals Calculates Confidence intervals	Yes	367	NC	NC	NC
[20]	Yes	Easy	No	Only for 50% estimate Considers population size Calculates Confidence intervals	No	382.6937	NC	NC	NC
[21]	No	Moderate	Yes (An excel file)	Provides No calculation Table is Only for 50% estimate	No	NC	NC	NC	NC
[22]	No	Easy	Yes	Provides only a monogram	No	383	NC	NC	NC
[23]	Yes	Easy	No	Multilingual (Persian, Arabic, French and English) Provides all 4 types of sample size calculations Calculates for non equal groups Provides calculations for correlation estimates	Yes	369	43	99+99	99+99

continuation of Table 2.

Reference	Providing Formulas	Usability	Downloadable File(s)	Type of Calculations Provided	Accurate calculation*	Common Sample Size Calculation Results			
						Estimation		Comparison	
						Qualitative <sup>1</sup>	Quantitative <sup>2</sup>	Qualitative <sup>3</sup>	Quantitative <sup>4</sup>
[24]	No	Easy	Yes	Provides PS: Power and Sample Size Calculation version 3.0 Software for download For windows, Macintosh and Linux. Calculates sample size and power	No	NC	NC	NC	28+28 Considering One SD
[25]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[26]	No	Advanced	Yes (A .jar file)	Calculates sample size for comparison of qualitative and quantitative outcomes and for survival data	Yes (just 3)	NC	NC	198 Total	504 Total Considering One SD
[27]	No	Moderate	No	Calculates sample size for usability studies.	NAP	NC	NC	NC	NC
[28]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[29]	Yes	Advanced	No	Just sample size calculation for two parallel samples in comparison of proportions Provides Non equal sample size calculation	No	NC	NC	10	NC
[30]	No	Easy	No	A survey sample size calculator Only for 50% estimate Considers population size Calculates Confidence intervals	No	383	NC	NC	NC
[31]	No	Advanced	No	Calculates power and Sample Size for Logistic Regression with Binary Covariates	NAP	NC	NC	NC	NC
[32]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[33]	No	Moderate	No	Provides all 4 types of sample size calculations	No	606 With Beta Error	69 With Beta Error	78+78	76+76
[34]	No	Advanced	Yes	Just provides links to other pages don't provide any direct calculation	-	NC	NC	NC	NC
[35]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[36]	No	Easy	No	Only for 50% estimate Considers population size Just for 95% and 99% confidence	No	383	NC	NC	NC
[37]	No	Easy	No	Multilingual (Chinese, English, French, Korean, Italian, Japanese, Spanish) Only for 50% estimate Considers population size Calculates Confidence intervals	No	383	NC	NC	NC

continuation of Table 2.

Reference	Providing Formulas	Usability	Downloadable File(s)	Type of Calculations Provided	Accurate calculation*	Common Sample Size Calculation Results			
						Estimation		Comparison	
						Qualitative <sup>1</sup>	Quantitative <sup>2</sup>	Qualitative <sup>3</sup>	Quantitative <sup>4</sup>
[38]	No	Easy	No	Only for 50% estimate Considers population size Calculates Confidence intervals	No	383	NC	NC	NC
[39]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[40]	Yes	Easy	Yes (Also online)	Only for 50% estimate Considers population size Calculates Confidence intervals	No	322	NC	NC	NC
[41]	No	Easy	No	Considers population size Calculates Confidence intervals	Yes	368	NC	NC	NC
[42]	No	Easy	No	Just for a qualitative estimation	Yes	368	NC	NC	NC
[43]	No	Easy	No	Considers population size	Yes	368	NC	NC	NC
[44]	Yes	Moderate	No	Provides calculations just for comparison of quantitative outcomes	Yes	NC	NC	NC	99+99
[45]	No	Advanced	No	Provides all 4 types of sample size calculations Calculates confidence intervals	No	1475	171	212+212	158+158
[46]	Yes	Advanced	No	Provides all 4 types of sample size calculations	No	1475.12	170.73	193.84 Total	223.25 Total Considering One SD
[47]	Yes	Easy	No	Considers population size	No	368	NC	NC	NC
[48]	No	Easy	No	Only for 50% estimate Considers population size	No	383	NC	NC	NC
[49]	Yes (Proportions) No (Complex Sample Survey)	Easy	No	Provides calculations for complex sample surveys Considers population size	Yes	368	NC	NC	NC
[50]	No	Advanced	No	Provides links to JavaScript based forms for simple power/sample size calculations (Comparing a Mean to a Known Value, Comparing Means for Two Independent Samples, Comparing a Proportion to a Known Value, Comparing Proportions for Two Independent Samples, Unmatched Case Control Studies)	No	NC	NC	97	112

continuation of Table 2.

Reference	Providing Formulas	Usability	Downloadable File(s)	Type of Calculations Provided	Accurate calculation*	Common Sample Size Calculation Results			
						Estimation		Comparison	
						Qualitative <sup>1</sup>	Quantitative <sup>2</sup>	Qualitative <sup>3</sup>	Quantitative <sup>4</sup>
[51]	No	Advanced	No	Only for comparison of 2 means Provides One and two sided calculations Considers single equal variances Calculates power	No	NC	NC	One sided=44 Two sided=56	NC
[52]	NR	NR	NR	Offers different software for sale Sample Size & Power Analysis (nQuery Advisor+nTerim) Missing Data Analysis (Solas) Analytical software (NCSS Statistical Analysis and Graphics Software)	NAP	NC	NC	NC	NC
[53]	No	Easy	Yes	Provides calculations for estimations	Yes (just 2)	367	42	NC	NC
[54]	No	Easy	No	Only for 50% estimate Considers population size Just for 95% and 99% confidence	No	383	NC	NC	NC
[55]	No	Advanced	No	Just calculates one sample for Non parametric survival	NAP	NC	NC	NC	NC
[56]	No	Easy	No	Considers response rate	Yes	368	NC	NC	NC
[57]	No	Moderate	Yes (An excel file)	Just calculates sample size for complex sample survey	NAP	NC	NC	NC	NC
[58]	No	Advanced	Yes (A setup file)	Calculates one tailed and two tailed Calculates for non equal groups Calculates for paired groups Calculates sample size for correlation estimates Considers default 95% power Provides power analysis and effect size calculations Provides calculations of P value according to common statistical tests Draws power and sample size plots	Yes (just 4)	NC	NC	97+97	100+100
[59]	No	Advanced	Yes	Provides sample size and power calculations Different equal and non equal groups Calculates one tailed and two tailed	Yes (just 3)	NC	NC	99+99	107+107

1=Sample size calculation for estimating a qualitative variable with expected 40% prevalence and 5% acceptable difference with 5% type I error.

2=Sample size calculation for estimating a quantitative variable with expected SD as 10 and 3 units acceptable difference with 5% type I error.

3=Sample size calculation for comparison of a qualitative variable with expected 40% and 60% prevalence in groups with 5% type I error and 80% power.

4=Sample size calculation for comparison of a quantitative variable with expected Mean±SD in groups as 25±7 and 22±8 with 5% type I error and 80% power.

\*In case of multiple calculations, all were not accurate, calculation numbers which was accurate (listed above) indicted in below parenthesis.



and some uses variances instead of standard deviations for calculations [16] this can cause potential errors if researchers use standard deviations instead.

Some other sites provided advanced calculations for complex studies, [12,26,49,55,57] in our study we did not rate and assess their calculations but when in simple ways of calculations there are so many faults how we can expect accurate results from those complex sites. Providing multi language support may be an ease of use for users with different languages other than English language. In our review we found two sites which provided multi language support [23,37].

Some good references for reviewing sample size calculation principles are used in this study and can be helpful for researchers [7,8]. Special statistical software like STATA, SAS and R also provides commands for sample size calculation and power analysis which can be useful for sophisticated approach.

## Conclusion

It seems that those sites which provide calculating formulas can be more trusted for online sample size calculation and at least there will be a way for researchers to check the calculation accuracy based on the provided formulas. Although researchers who perform sophisticated studies need special assistance of statisticians for sample size calculation in their studies.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

Authors' contributions	AM	FT	MMV	SYK	SEF	MA
Research concept and design	✓	✓	--	--	--	--
Collection and/or assembly of data	✓	✓	✓	✓	✓	--
Data analysis and interpretation	✓	✓	✓	✓	✓	✓
Writing the article	✓	✓	✓	✓	✓	✓
Critical revision of the article	✓	--	--	--	--	✓
Final approval of article	✓	--	--	--	--	--
Statistical analysis	✓	✓	✓	✓	--	--

## Acknowledgement

The authors thanks to Tehran University of Medical Sciences for providing the facilities to run an online sample size calculator and for supporting the study.

## Publication history

Editor: Jimmy Efirid, East Carolina University, USA.  
 Received: 08-Jan-2013 Final Revised: 24-Feb-2014  
 Accepted: 26-Feb-2014 Published: 14-Apr-2014

## References

- Albert NM and O'Connor PC. **How many subjects do I need in my research sample?** *Clin Nurse Spec.* 2012; **26**:302-4. | [Article](#) | [PubMed](#)
- Schlesselman JJ. **Sample size requirements in cohort and case-control studies of disease.** *Am J Epidemiol.* 1974; **99**:381-4. | [Article](#) | [PubMed](#)
- Lwanga SK and Lemeshow S. **Sample Size Determination in Health Studies: A Practical Manual,** World Health Organization (1991). | [Pdf](#)
- Mace AE. **Sample-size determination.** Reinhold, New York, 1964. | [Book](#)
- Mellor B. **Sample size methodology.** M. M. Desu and D. Raghavarao, Academic Press, Boston, 1990. *Statist Med.* 1992; **11**:562–563. | [Article](#)
- Ahlbom A. **Biostatistics for Epidemiologists.** Lewis Publishers, Boca Raton, FL. 1993. | [Website](#)
- Meysamie A and Peysepar R. **Medical Statistics.** Golbarg Press. 2012 (Persian).
- Ryan TP. **Sample size determination and power.** John Wiley & sons, USA. 1945. | [Website](#)
- Donner A. **Approaches to sample size estimation in the design of clinical trials--a review.** *Stat Med.* 1984; **3**:199-214. | [Article](#) | [PubMed](#)
- 6sixsigma. | [Website](#)
- Black berry world. | [Website](#)
- Sample Size for Microarray Experiment. | [Website](#)
- BOLD Educational Software Sample Size Calculator. | [Website](#)
- DusickCalculating Sample Size. | [Website](#)
- Easycalculation. | [Website](#)
- Epi Tools (Ausvet animal health services). | [Website](#)
- Statistical considerations for clinical trials and scientific experiments. | [Website](#)
- Java applets for power and sample size. | [Website](#)
- Advanced Sample Size Calculator. | [Website](#)
- Tamaño de la muestra la calculadora. | [Website](#)
- The research advisor. | [Website](#)
- SampleSize The Research Advisors 2006. | [Website](#)
- Online Sample Size Calculator. | [Website](#)
- Power and Sample Size Calculation. | [Website](#)
- Answers Research. | [Website](#)
- Statistical Programs for Clinical Trials. | [Website](#)
- Usability Sample Size Calculator. | [Website](#)
- Sample Size Calculator (Calculator.net). | [Website](#)
- Sample Size Calculator: Two Parallel-Sample Proportions. | [Website](#)
- Calculate a survey sample size (number of respondents needed). | [Website](#)
- Power/Sample Size Calculation for Logistic Regression with Binary Covariate(s). | [Website](#)
- Evaluate your sample size using Decision Point's sample size calculator. (*Decision point*). | [Website](#)
- Researcher's Toolkit, (DSS Research). | [Website](#)
- Power and Sample Size Programs. | [Website](#)
- Florida law firm. | [Website](#)
- Sample Size Calculator. | [Website](#)
- Solutions: Resources | Sample Size Calculator. | [Website](#)
- IBP Strategy and Research. | [Website](#)
- LCWA. | [Website](#)
- MaCorr Research Solutions. | [Website](#)
- National Statistical Services. | [Website](#)
- Openuptomail. | [Website](#)
- Pivotal Research Inc. | [Website](#)
- SISA (simple interactive statistical analysis). | [Website](#)
- Ques gen system inc. | [Website](#)
- Sample Size Estimation. | [Website](#)
- Raosoft Inc. | [Website](#)
- Sample Size Calculator. | [Website](#)
- Kevin's web page. | [Website](#)

- 50. UCSF Biostatistics: Power and Sample Size Programs | [Website](#)
- 51. Power & Sample Size Calculator. | [Website](#)
- 52. nQuery Advisor + nTerim. | [Website](#)
- 53. StatPac Statistics Calculator Sample Size Estimation. | [Website](#)
- 54. Creative research system. | [Website](#)
- 55. SWOG: Southwest Oncology Group. | [Website](#)
- 56. Torque Data. | [Website](#)
- 57. STEPS Sample Size Calculator and Sampling Spreadsheet, WHO. | [Website](#)
- 58. G\*Power: Statistical Power Analyses for Windows and Mac. | [Website](#)
- 59. STATA. | [Website](#)

**Citation:**

Meysamie A, Taei F, Mohammadi-Vajari M-A, Yoosefi-Khanghah S, Emamzadeh-Fard S and Abbassi M. **Sample size calculation on web, can we rely on the results?** *J Med Stat Inform.* 2014; **2**:3.  
<http://dx.doi.org/10.7243/2053-7662-2-3>