

Obtaining the "sample"

Ogunbanjo GA

In research, the sample is a selected subset of the study population. It may be randomly or non-randomly selected, representative or non-representative of the study population. To have a clear understanding on how to obtain the sample, it is important to have a good idea of various types of sample. These include:

- **Cluster sample:** In this type of sample, the unit selected is always a group of persons rather than an individual. An example of cluster sample may be a group of children in a household.
- **Sample of convenience:** As this suggests, it is sample selected by easily employed but basically non-probabilistic methods. The opinion poll "man-in-the-street" surveys fall into this category. The sample of convenience is also referred to as the "Grab" sample.
- **Random sample:** This is the commonest type of sample used in quantitative research. In a random sample, each individual has an equal chance of being included in the study. Random sample can be *simple*, that is, each individual is assigned a number and using a table of random numbers the sample is selected, or *stratified* in which the study population is divided into distinct subgroups (strata) according to some important characteristics with the random sample selected from each subgroup.
- **Systematic sample:** This is a procedure of selection based on some simple, systematic rule such as every second or third person available or patients with odd numbers as the last digits of their medical files etc. The danger of this type of sample is that, it may lead to errors that invalidate generalizations.
- **Purposive sample:** This sample is often used in *qualitative* research, in which the sample is composed of certain elements which contain the most characteristic, representative or typical attributes of the population.¹ This type of sample is based entirely on the judgement of the researcher.
- **Snowball sample:** This involves approaching a single individual who is

involved in the phenomenon under investigation, to gain information on other similar persons. The latter is in turn requested to identify other people who may make up the sample. In this way the researcher proceeds until a sufficient number of participants make up the sample. This technique is excellent when investigating a relatively unknown or rare phenomenon.

It is obvious that the most reliable type of sample is the "random sample" in that it gives every individual a chance of being included in the study. By now, you may be wondering how to determine the sample of your next study. The first step would be to define your **study population**. The ideal situation would be for the study population to be the same as the sample, in order to include everyone. But if the study population is large, the large sample becomes unmanageable in terms of *time* and *cost*. On the other hand if the sample were too small, the results obtained would not be useful or conclusive. But the larger the sample, the more sure you can be that the responses obtained from the study truly reflect the population. The latter relates to the *confidence level* of the study, that is, the larger the sample size, the smaller the confidence interval.

There are formulae and computer statistical programs available that perform

sample size calculations e.g. Epi-Info, which is a freeware on the Internet. Most of them assume that simple random sampling would be used. Stoker offers an indication of what the size of a sample ought to be (Table I).² This table provides a simple, but accurate estimation of the sample based on various study populations and can be used in the absence of a computer statistical software or formula. I hope that this article has provided the family practitioner with some understanding of the various types of sample and a simple table to calculate the sample of a quantitative study. The approach for qualitative studies is different and will be discussed in a future article.

References

1. Singleton R, Straits BC, Straits MM, McAllister RJ. *Approaches to social research*. New York: Oxford University Press: 1988.
2. Stoker DJ. *Sampling*. In: de Vos AS, editor. *Research at grass roots*. 1st ed. Pretoria: JL van Schaik Publishers; 1998: 192-93.

Prof. Gboyega A Ogunbanjo
MBBS, MFGP (SA), M Fam Med (Medunsa),
FACRRM
Dept. of Family Medicine & Primary Health
Care,
Medunsa, Pretoria.

Table I: Guideline for sampling²

Study population	Percentage suggested	Number of respondents
20	100%	20
30	80%	24
50	64%	32
100	45%	45
200	32%	64
500	20%	100
1 000	14%	140
10 000	4,5%	450
100 000	2%	2 000
200 000	1%	2 000