Selecting the sample

Allison Shorten,[1] Calvin Moorley[2]

10.1136/eb-2014-101747

[1]Yale University School of Nursing, West Haven, Connecticut, USA
[2]University of East London, School of Health, Sport and Bioscience, Stratford, London, UK

Correspondence to: **Dr Allison Shorten**, Yale University School of Nursing, PO Box 27399, West Haven, CT 06516-7399, USA; allison.shorten@yale.edu

Introduction

Sample selection is a key factor in research design and can determine whether research questions will be answered before the study has even begun. Good sample selection and appropriate sample size strengthen a study, protecting valuable time, money and resources. In the context of healthcare research, poor design could lead to use of harmful practices, delays in new treatment and lost opportunities for high quality care. Not every study can achieve design perfection and researchers often seek a balance between the ideal sample and one that is feasible or convenient, acknowledging the limitations of their design decisions. Sample selection is far from simple but here are some of the techniques to think about as you read research and make the most out of your research endeavours.

Sampling the population

It is critical to take the time to clearly identify the population of interest for the specific research question. Nursing researchers are usually interested in answering questions about very specific patient populations which can span an incredible array of possibilities applying to international, national, local and organisational contexts. Research populations closely reflect nursing specialties, some of which are gender (eg, pregnant women) and age specific (eg, adolescent diabetes). It is rarely feasible to conduct a study that reaches every patient in the population of interest, therefore a subset or sample of that population is selected for study.

Different sampling methods are used depending on the aim of the study and whether the research question seeks a confident answer about the population of interest. If it does then the sample/s should represent the population for inferences to be made. Not all research questions depend on making inferences and there are many examples in qualitative research where the aim is theory development or exploration of patient experiences and inferences are not the focus.

Sampling methods: quantitative research

Probability (representative) sampling includes techniques used to select a sample that clearly represents a specific population. Participants are randomly selected so that all members of the population have an equal chance of being selected in the sample, eliminating the possibility of sample selection bias.

Example: Let's say Hospital A is interested in the opinions of all 1000 of their patients who underwent knee surgery during the calendar year 2013 (the population). The Quality Assurance (QA) budget will not allow for *all* patients to be surveyed but they want to be sure the sample represents all patients. One way to do this is by selecting a random sample of a prespecified number of patients.

If *simple random selection* was used, the desired number of patients (eg, 200 or 20%) would be randomly selected from a complete list of 1000 patients using random numbers between 1 and 1000. Random numbers can be generated using a spreadsheet program or an online randomiser that can export a set of numbers into spreadsheet programs such as excel http://www.randomizer.org/form.htm[1]

A *stratified random sample* could be selected for a more targeted approach. If the hospital was aware that there were only 400 women who had knee surgery in that year and wanted the sample to represent known gender proportions for all patients having knee surgery they could stratify the populations first. In this case, the patient population would be divided (stratified) according to gender before simple random selection was used to identify 40% (n=80) female and 60% (n=120) male patients for the sample.

Non-probability sampling (non-representative) although usually systematic and capable of answering research questions, can suffer from issues associated with sample selection bias. Let's say the QA nurse at Hospital A is interested in patient opinions about the new hospital admission and discharge process for knee surgery patients. Owing to budget limitations and the fact that the QA nurse only works 1 day/week, the nurse uses a convenient *purposive sample* to survey patients who experience knee surgery every Friday when she is in the hospital. This particular convenience sample may be representative of patients who experienced knee surgery on Fridays (the sample); however, it would not necessarily represent all patients who experienced knee surgery in that hospital (the population). There may be differences in the characteristics of patients having surgery on Fridays. Perhaps the surgeons who operate on Fridays mostly see sports-injured male patients (age range 25–35 years) with private health insurance compared with the surgeons who operate during the remainder of the week serving male and female publicly insured patients (age range 35–55 years). Potential differences in sample characteristics such as age, gender, level of presurgery fitness, income, education and employment could be unique to the sample and fail to answer the question for all patients in the hospital.

Sample size

Selecting the right sample size is about predicting in advance that the sample size will be large enough to give adequate 'power' to the study. The 'power' of a study can be defined as the probability of correctly identifying that the intervention produces a treatment effect if one actually exists. Statisticians state that we correctly reject the null hypothesis of no treatment effect, or alternatively as the probability of avoiding a type II error.

There are many sample size power calculators available online such as http://www.stat.ubc.ca/~rollin/stats/ssize/index.html[2]

To calculate a sample size the researcher needs to understand the type of variables being compared (eg, mean values vs proportions or percentages); be able to specify the variables of interest for each sample; estimate the likely size of the treatment effect: estimate the SD of the combined samples (if comparing means); specify whether the calculation should be based on a one-sided or two-sided test (eg, if it is an intervention study and it is thought that the treatment may have the potential to either improve or harm patient outcomes, then a two-sided test should be used); and finally to specify values for both 'α' and 'power'. The 'α' reflects our level of willingness to tolerate a type I error—incorrectly rejecting a null hypothesis (a false positive). Conventionally, researchers set α at 0.05 (5%) when conducting sample size power calculations. The 'power' indicates the probability that the samples selected will detect an intervention effect if one exists. There is a conventional value of 0.8 (or 80%).

Sample size calculation is not an exact science and sometimes decisions are made without prior research to inform the estimates. Pilot studies are often needed to provide data to guide sample size predictions. Watch out for studies where sample size is small with unrealistic or unsubstantiated estimates for large intervention effects. This can be a sign that the smaller sample size has been decided first with post hoc power calculations to justify this.

Sampling methods: qualitative research

Similar to quantitative studies, qualitative research questions determine the sample and sampling strategies used. When we ask 'why' and 'how' questions to illuminate, unravel and understand complex human psychosocial issues, the focus is not on predetermined hypotheses or generalisable results. Samples may be sought for focus groups, personal interviews (structured, semistructured or unstructured) or for participatory, observational and ethnographic work.[3] This type of research often involves a small number of participants to explore in-depth feelings, experiences and held beliefs. A sample can be as small as one person (a case study) and sampling ends when saturation of emerging themes has been reached. Data analysis begins at data collection, requiring a design that is flexible and iterative, as the researcher may move backward to move forward in response to emerging themes.

There are three broad qualitative sampling approaches; convenience, judgement and theoretical sampling.

Convenience samples are recruited based on accessibility to the researcher. An example is snowball sampling where participants guide the researcher to the next participant, via acquaintances or social networks. Snowballing is useful when studying hard to reach groups but can result in a sample bias. Convenience sampling is the least rigorous technique used and can result in poor-quality data, which should be balanced against resource savings (time and money) or convenience.

Judgement sampling, commonly referred to as purposeful or selective sampling, relies on the judgement and practical knowledge of the researcher to identify and select participants. A framework for selection can be developed from variables identified in the literature, combined with practical knowledge of the phenomena. This is a more systematic strategy and can increase sample credibility using a wide range of participants, for example, those with in-depth experience or special knowledge of the research topic.

Theoretical sampling is a hallmark of the grounded theory approach.[4] Here, the sample is theory driven and the researcher continues data collection and analysis until theoretical saturation is reached. The samples are used to build and produce an interpretative theory. This is useful when the aim is theory and concept development, grounded in or emergent from real-world life events and circumstances.

Conclusion

Sample selection is far from simple so check your research design and sample selection plan with an expert statistician or research consultant early, and before the data collection begins. Designing a study or writing research questions to suit a sample that is merely convenient to reach is dangerous and readers should always be alert to signs of poor practice in sample selection. Great samples are worth the effort.

Competing interests None.

References

1. **Urbaniak GC,** Plous S. Research Randomizer Social Psychology Network. http://www.randomizer.org/form.htm
2. **UCSF Department of Epidemiology and Biostatistics.** Power and Sample Size Calculations. http://www.stat.ubc.ca/~rollin/stats/ssize/index.html
3. **Green J,** Thorogood N. *Qualitative methods for health research.* 2nd edn. London: Sage, 2009.
4. **Glaser B,** Strauss A. *The discovery of grounded theory: strategies for qualitative research.* Chicago: Aldine, 1967.