

If you could just provide me with a sample: examining sampling in qualitative and quantitative research papers

When undertaking any research study, researchers must choose their sample carefully to minimise bias. This paper highlights why practitioners need to pay attention to issues of sampling when appraising research, and discusses sampling characteristics we should look for in quantitative and qualitative studies. Because of space restrictions, this editorial focuses on the randomised controlled trial (RCT) as an example of quantitative research, and grounded theory as an example of qualitative research. Although these 2 designs are used as examples, the general principles as outlined can be applied to all quantitative and qualitative research designs.

What is sampling?

Research studies usually focus on a defined group of people, such as ventilated patients or the parents of chronically ill children. The group of people in a study is referred to as the *sample*. Because it is too expensive and impractical to include the total population in a research study, the ideal study sample represents the total population from which the sample was drawn (eg, all ventilated patients or all parents of chronically ill children). This point—that studying an *entire* population is, in most cases, unnecessary—is the key to the theory of *sampling*. Sampling means simply studying a proportion of the population rather than the whole. The results of a study that has assembled its sample appropriately can be more confidently applied to the population from which the sample came. Using the examples of samples provided at the start of the paragraph, we can see that Chlan *sampled* 54 patients from a *population* of patients who required mechanical ventilation,¹(see *Evidence-Based Nursing* 1999 April, p49) whereas Burke *et al* *sampled* 50 children (and their parents) from a *population* of all children requiring admission to hospital for chronic health conditions.²(see *Evidence-Based Nursing* 1998 July, p79) In both studies the researchers wanted to say something that would apply to the population by examining a small portion of those populations.

When reading a paper that sets out to say something about a population by studying a sample, readers need to assess the *external validity*. External validity is the degree to which the findings of a study can be generalised beyond the sample used in the study. The ability to generalise is almost totally dependent on the adequacy of the sampling process. Nurses should consider several possible threats to external validity when appraising a paper and deciding whether the results could be applied to patients in their care:

- *Unique sample selection*: study findings may be applicable only to the group studied. For example, the findings of a study of a telephone support programme for caregivers of people with Alzheimer's disease that recruited a sample from local Alzheimer's Association branches would not necessarily apply to the population of caregivers as a whole because most caregivers do not belong to such local groups, and therefore are different from the sample.

- *Unique research settings*: the particular context in which the study takes place can greatly affect the external validity of the findings. For example, Tourigny sampled 6 African-American youths to learn about deliberate exposure to HIV.³ (see *Evidence-Based Nursing* 1998 October, p130) The study took place in the context of extreme poverty in a uniquely deprived US urban setting. The opportunity for a suburban district nurse in the UK to apply the theory generated by this study may be limited by the very different social settings involved.
- *History*: the passage of time can affect the findings. For example, studies of mechanisms for implementing research findings within healthcare organisations might be affected by the organisational and structural changes that occur at local and national levels over time. Examples of such reorganisation include the effects of separating healthcare provision from purchasing through the NHS "internal market" of the late 1980s and the creation of health maintenance organisations in the US.
- *Unique research constructs*: the particular constructs, concepts, or phenomena studied may be specific to the group sampled. For example, researchers evaluating the concept of "quality" in healthcare should recognise that professionals and consumers may differ in their perceptions of the concept and should be explicit about how they actually measured it.

Sampling in quantitative research

Quantitative research is most often used when researchers wish to make a statement about the chance (or probability) of something happening in a population. For example, a person is 65% less likely to die a cardiac related death if she eats a Mediterranean type diet compared with the recommended Step 1 diet of the American Heart Association.⁴ (See *Evidence-Based Nursing* 1999 April, p48.) Quantitative studies usually use sampling techniques based on *probability theory*. *Probability sampling*, as it is known, has 2 central features:

- The researcher has (in theory) access to all members of a population
- Every member of the population has an equal and non-zero chance of being selected for the study sample. In other words they cannot have "no chance" of being sampled. Three concepts relevant to probability samples are *sampling error*, *random sampling*, and *sampling bias*. Each of these will be described.

SAMPLING ERROR

Probability samples allow researchers to minimise *sampling error* in that they give the highest chance of the sample being representative of the total population. *Sampling error* occurs in all probability samples and is unavoidable because no sample can ever totally represent the population. There will always be a gap between a sample's representativeness and

the population's known or unknown characteristics—the *sampling error*. Readers of quantitative research should look for evidence that the researchers tried to combat sampling error. Specifically, the authors should identify the study sample using a random selection process and should provide substantiation for the sample size. The size of the sampling error generally decreases as the size of the sample increases.

RANDOM SAMPLING

Random selection works because as individuals enter the sample, their characteristics (which are different from the population) balance the characteristics of other individuals. For example, in a randomly selected sample of users of mental health services, there will be users who are from upper socioeconomic groups, and these will be balanced by those who are from lower socioeconomic groups.

Successful random sampling requires a sufficiently large sample. If the sample is large enough, then differences in outcomes that exist between groups will be detected statistically, whereas if it is too small, important differences may be missed. One of the clues that can alert the reader to a study that is not large enough is the confidence interval around a study finding. Although not all studies provide confidence intervals, these are becoming increasingly popular in the reporting of quantitative studies. A confidence interval provides a statement on the level of confidence that the true value for a population lies within a specified range of values. A 95% confidence interval can be described as follows: "if sampling is repeated indefinitely, each sample leading to a new confidence interval, then in 95% of the samples, the interval will cover the true population value.⁵ The larger the sample size, the more narrow the confidence interval, and therefore the more precise the study finding. In the study by Egerman *et al* in this issue of *Evidence-Based Nursing* (p73), the relative risk of necrotising enterocolitis with oral versus intramuscular dexamethasone was 5.1 with a 95% confidence interval of 0.8 to 36.6. This range means that the true relative risk could be as low as 0.8 or as high as 36.6 and because a relative risk of 1.0 (meaning no difference between groups) is included in this range, the conclusion is that there is no difference between the 2 methods of dexamethasone administration in the risk of necrotising enterocolitis. The reader should, however, take careful note that this confidence interval is very wide. It is possible that with a larger sample size and, consequently, a narrower confidence interval, a statistically significant difference between groups may have been found because the confidence interval may no longer include the relative risk of 1.0.

SAMPLING BIAS

Unlike sampling error, which cannot be avoided completely, *sampling bias* is usually the result of a flaw in the research process. It is systematic, and increasing the size of the sample just increases the effect of the bias. Sampling bias occurs when the sample is not representative of the population. An example of *sampling bias* has already been highlighted in the section on external validity: sampling only caregivers from Alzheimer's Association groups introduces a sampling bias if the study aim is to generalise to the whole population of caregivers of people with Alzheimer's disease. Bias, in the context of RCTs, can be thought of as "...any factor or process that tends to deviate the results or conclusions of a trial away from the truth."⁶ Two important biases related to sampling that could affect generalisability of study findings are referral filter bias and volunteer bias. In referral filter bias, the

selection that occurs at each stage in the referral process from primary to secondary to tertiary care can generate patient samples that are very different from one another.⁷ For example, the results of a study of patients with asthma under the care of specialists are not likely to be generalisable to patients with asthma in primary care settings. In volunteer bias, people who volunteer to participate in a study may have exposures or outcomes (eg, they tend to be healthier) that differ from those of non-volunteers.⁸

When appraising a research report, readers can ask a simple set of questions to assess whether sampling bias exists:

- *Who* was included in the sample?
- *What* was the *source* of recruitment into the study? Community or hospital? Specialised centres or general hospitals?
- *How* were patients *recruited*? Approached by healthcare professionals or study researchers? Paid or unpaid?
- *Which* patients were *approached* to be in the study? Consecutive patients? Volunteers or only people the researcher thought looked like useful candidates?
- *What* were the *inclusion and exclusion criteria* for the sample? What were the demographics (eg, age and sex) of the sample? Did they have any medical conditions or a history of previous interventions that may have a bearing on the results of the study?

Some research designs—specifically, RCTs—use probability theory slightly differently. In these studies, individuals might be initially *sampled* on a non-probability basis (eg, *all* people attending an asthma clinic in 1 year) and then *randomly allocated* to an experimental or control group (eg, self management education or usual care). It is in the process of allocation that probability theory comes into play. Through random allocation, and provided that the sample is large enough, all known and unknown confounders are equally distributed among the groups and therefore, at the end of the study, any differences among the groups can be attributed to the intervention. Randomisation *alone*, however, does not allow the researcher to generalise to a population. Questions of sample size and the representativeness of the sample as a whole are still important. Readers should look for evidence that the sample was large enough and that there were no baseline differences between the experimental and control groups. If they are satisfied on these 2 counts, then they can go on to consider the similarity between the study sample and their own patients. If they are similar, then the study sample can be said to be *representative* of their patients, and readers can feel confident in generalising the study findings.

Sampling in qualitative research

Not all research designs are concerned with generalising from a sample to a population of people. Qualitative studies use rich and deep description to inform our understanding of concepts and contribute to broader theoretical understanding. For example, Wilson's study of the experiences of older adults making the transition to life in a nursing home says nothing about the *amount* of experience or the *probability* of "adjustment" to a nursing home; instead, the author simply describes what that experience might be like.⁹ (See *Evidence-Based Nursing* 1998 Jul, p96.) On the basis of this description, the author proposes a 3 stage theory about the process of adjustment to a nursing home placement. What this means is that nurses should consider the possibility that these stages might need to be addressed with patients undergoing this transition. Good qualitative enquiry leads to theoretical generalisability rather than statistical inference.

Because we want to generalise about the shape or content of a concept or contribute to a theoretical understanding of an aspect of healthcare, sampling in qualitative research is driven by a very different set of concerns than those of quantitative designs.

“The purpose [of sampling in qualitative research] is not to establish a random or representative sample drawn from a population but rather to identify specific groups of people who either possess characteristics or live in circumstances relevant to the social phenomenon being studied. Informants are identified because they will enable exploration of a particular aspect of behaviour relevant to the research. This allows the researcher to include a wide range of types of informants and also to select key informants with access to important sources of knowledge.”¹⁰

Qualitative researchers therefore use non-probability sampling techniques as the basis for their studies. For example, if a researcher wanted to know more about the content of unpleasant experiences among people receiving primary health care, it would be of no value to ask a random sample of the population because most people have relatively pleasant experiences of consuming primary care. It would be far more advantageous to interview people who were unhappy with the care they received (perhaps identified from complainant records). If we wanted to know the extent of unhappy experiences among consumers then, of course, the sampling strategies would be reversed and we might randomly select individuals from the whole population and attempt to measure the amount of dissatisfaction.

NON-PROBABILITY SAMPLING

Just as there are various approaches to sampling in quantitative studies, there are several approaches to non-probability sampling used in qualitative research. The same questions asked of probability samples can be asked of non-probability samples. Questions of sample size, however, are not as important here. Rather, it is the “fitness for purpose”—or quality—of the sample that should be described in the paper. This is difficult to judge in qualitative studies because the adequacy of the sample can often only be judged by the quality of the analysis of the data generated. In some qualitative studies, sampling continues until no new themes emerge from the data, a point called data saturation. The earlier questions remain useful as a way of focusing your scrutiny.

- *Who* was included in the sample? That is, are the participants able to adequately contribute to knowledge and understanding?
- *What* was the *source* of recruitment into the study? Is this context likely to influence the shape of the analysis presented?
- *How* were patients *recruited*? Approached by healthcare professionals or study researchers? Paid or unpaid? Are these factors likely to alter the stories, accounts, or behaviours presented, and if so how did the researcher account for this?
- *Which* patients were *approached* to be in the study? Consecutive patients? Volunteers or only people the researcher thought looked like useful candidates?
- *What* were the *inclusion and exclusion criteria* for the sample? For example, what were the demographics (eg, age and sex) of the study sample? Did patients have any medical conditions or history of previous interventions that may have a bearing on the results of the study—in qualitative research this usually means the accounts people provide or behaviours presented.
- All of these questions remain relevant, but clues that should help readers to assess the adequacy of a qualitative

sample include the believability of the final product and the sense of completeness in the analysis presented. A solid qualitative paper should generate theory or description that encompasses a range of experiences or values and makes explicit the limitations of qualitative research in regard to conventional generalisability.

Qualitative and quantitative sampling: a final word of caution

This overview has used the qualitative-quantitative distinction to highlight the primary differences between probability and non-probability sampling strategies. In reality, however, this distinction is often blurred and many studies use a combination of 2 or more approaches. This is not necessarily a criticism—sampling is an integral part of the application of research methodology and because of this it should be guided by the research question. Like so much in evidence-based health care, the key questions to ask about any study’s sampling strategy relate to the “fitness for purpose.” Three questions that capture the quality of any sample are:

- *Are the patients in the study similar enough to my patients?* In quantitative terms, this enables you to judge whether the results of the research can be applied to your population. In qualitative terms, the similarity of patients enables you to judge whether the concept or experiences being explored would be meaningful to your practice.
- *Was the sample selected in such a way that it could introduce bias into the research*—this is as applicable to qualitative research as it is to quantitative approaches. In qualitative research, however, the boundaries or limitations of theoretical generalisability are usually encompassed in the description of the participants.
- *Was the sample large enough?* In quantitative research, this means “was it large enough to detect a difference between groups if a difference existed?” In qualitative research, this means “were there enough people (or enough time spent with a few people) to provide rich meaningful description and convincing analysis?”

In summary, an understanding of how patients are assembled for a study, plus additional information about their characteristics (eg, age, sex, and disease severity), will help readers to decide whether the study sample resembles their own patients enough for them to apply the results to their own practice.⁷

CARL THOMPSON, RN, PhD

*Centre for Evidence Based Nursing,
Department of Health Studies,
University of York,
York, UK*

- 1 Chlan L. Effectiveness of a music therapy intervention on relaxation and anxiety for patients receiving ventilatory assistance. *Heart Lung* 1998;**27**:169-76.
- 2 Burke SO, Handley-Derry MH, Costello EA, *et al*. Stress-point intervention for parents of repeatedly hospitalized children with chronic conditions. *Res Nurs Health* 1997;**20**:475-85.
- 3 Tourigny SC. Some new dying trick: African American youths “choosing” HIV/AIDS. *Qual Health Res* 1998;**8**:149-67.
- 4 de Lorgeril M, Salen P, Martin JL, *et al*. Mediterranean dietary pattern in a randomized trial: prolonged survival and possible reduced cancer rate. *Arch Intern Med* 1998;**158**:1181-7.
- 5 Satin A, Shastry W. *Survey sampling: a non-mathematical guide*. Ottawa: Statistics Canada, 1985: 7-32.
- 6 Jadad A. *Randomised controlled trials*. London: BMJ Publishing Group, 1998:28.
- 7 Sackett DL, Haynes RB, Guyatt GH, *et al*. *Clinical epidemiology. A basic science for clinical medicine*. Second edition. Boston: Little, Brown and Company, 1991.
- 8 Sackett DL. Bias in analytic research. *J Chron Dis* 1979;**32**:51-63.
- 9 Wilson SA. The transition to nursing home life: a comparison of planned and unplanned admissions. *J Adv Nurs* 1997;**26**:864-71.
- 10 Mays N, Pope C. *Qualitative research in health care*. London: BMJ Publishing Group, 1996:12-3.