Journal of Orthopaedic & Sports Physical Therapy
2001;31(12):753–758

SPECIAL REPORT

A Clinician's Guide to Specification and Sampling

Kathryn E. Roach, PhD, PT[1]

This article will describe and discuss the implications of various steps in the process of selecting a sample for a research study and should assist clinicians in deciding whether and how to apply specific research findings to clinical care. *J Orthop Sports Phys Ther 2001;31: 753–758.*

Key Words: *nonprobability sampling, probability sampling, specification*

Evidence-based practice is possible only if clinicians are willing to use findings from research studies to make clinical decisions concerning the individual patients in our care.[9] Applying research findings from the group of subjects used in a study (the sample) to patients with similar characteristics (the target population) in the clinic is a multistep process. Initially, clinicians must decide whether their patients are comparable to the type of patients described in the study.[1,6] The clinician's patient must be sufficiently similar to the subjects in the research study for the research to be relevant to clinical decisions. Second, the clinician must decide whether the process used to select study participants produced a sample of subjects that was representative of the target population.[1,6] Finally, the data must be interpreted to decide whether conclusions can be reached about the target population based on information contained in the sample.[1] This final step uses statistical inference, which is based on the assumption that the study sample is a randomly selected subset of the target population. This assumption is often violated in clinical research studies,[3] and clinicians must therefore decide how closely the sample used in the study approximates a random sample.

SPECIFICATION

A population is a group of persons, objects, or events that meet a specified set of criteria.[1,2,4,6] Specification is the process by which a researcher defines the population being studied[3,4] in terms of inclusion and exclusion criteria. Typically, 2 types of populations are specified when describing a research study, the target population and the accessible population.[4,7]

[1] *Associate professor and associate director for research, University of Miami School of Medicine, Department of Orthopaedics and Rehabilitation, 5915 Ponce de Leon Boulevard, 5th Floor, Coral Gables, FL 33146-2406. E-mail: keroach@miami.edu*

Inclusion Criteria for the Target Population

A target population is the overall group to which the researcher hopes to generalize the findings of the study. The target population is specified by listing the inclusion criteria, that is, the characteristics of patients who are eligible to be included in the study (Table 1).[3,4,7] Because the results of studies should be considered applicable only to the same types of patients as those studied, clinicians seeking to use a research study as the basis for a clinical decision should examine the inclusion criteria carefully. For example, the findings of a study including only female subjects over the age of 65 may not be particularly useful in making clinical decisions about men between the ages of 18 and 25. Similarly, the findings of a study conducted using subjects with mild Parkinson's disease may not be very helpful to a clinician treating individuals with advanced Parkinson's disease.

Exclusion Criteria

Having specified those subjects who are eligible to participate in a research study (inclusion criteria), the researcher must also decide what subject characteristics would require the subject to be excluded from the study (Table 1).[3,4,7] Some subjects must be excluded in or-

TABLE 1. Specification.

Inclusion Criteria		
Target Population	**Accessible Population**	**Exclusion Criteria**
Age	Time period	Factors interfering with data quality
Gender	Geographic location	Factors interfering with interpretation
Ethnic group		Factors contraindicating intervention
Clinical characteristics		Ethical considerations
Disease severity		Unwilling to participate

der to preserve the credibility of the study. Unfortunately, researchers must often decide on a compromise between study feasibility and study generalizability.[3,4,7] For example, if one of the primary outcome measures for a research study requires that the subjects follow complex directions to perform a physical task, subjects who cannot follow such directions because of cognitive problems must be eliminated from the study. If a planned exercise intervention might be dangerous for subjects with brittle diabetes, then such subjects would need to be dismissed. Subjects may also be excluded for ethical reasons; for instance, pregnant women are often excluded because of possible risk to the unborn child.

Exclusion criteria are also important to the researcher in order to ensure that the findings of the study are valid. Subjects with characteristics that would interfere with interpreting the study findings must be excluded. Exclusion criteria are also important to a clinician attempting to utilize the findings of a research article to inform clinical practice. If the design of the study required the researcher to eliminate subjects who have characteristics similar to those under the clinician's care, the value of the study's findings may be substantially reduced. For example, a therapist working with patients with cognitive problems might find it difficult to utilize the findings of a study in which subjects with cognitive problems were excluded. Again, it is up to the clinician to decide to what extent the exclusion criteria have diminished the value of the study for informing clinical practice.

Inclusion Criteria for the Accessible Population

Although it would be ideal to draw a sample directly from the entire target population, in most cases this is impossible to do. Pragmatically, research must be performed on subjects in some specific location and during a designated period of time. The accessible population is the portion of the target population that is available to a particular researcher for study.[4,7] The accessible population is specified by stating inclusion criteria dealing with geographic location and time frame.[4,7] It is from this accessible population that the sample of subjects is actually drawn. There are 2 primary types of accessible populations used in clinical research, clinic-based and community-based.[4]

Clinic-based accessible populations Clinic-based accessible populations are defined by the geographic location of the clinical facility and the time frame during which subjects are recruited. Compared to community-based accessible populations, clinic-based accessible populations have the advantage of containing a relatively large proportion of potential subjects with the desired clinical characteristics.[4] For this reason, it is usually much easier and less expensive to select subjects from a clinic-based accessible population. Unfortunately, subjects selected from a clinic-based population may differ from the target population in important and predictable ways.

First, characteristics of the clinic may greatly influence the attributes of the patients treated there. For example, patients treated in a tertiary care medical center have often been examined and referred by physicians practicing in smaller community hospitals. Therefore, patients referred to tertiary care centers may tend to have more severe or complicated disease of longer duration. A sample drawn from such a center would produce a distorted picture of the features and prognosis of the disease being studied.

A second problem with clinic-based samples is that not all individuals with problems such as low back pain seek medical attention. Subjects selected from a clinic population may differ from subjects selected from a community sample on those psychosocial, economic, and disease severity factors that are related to seeking and receiving medical attention for their condition. This type of problem is lessened for conditions, such as hip fracture, for which inpatient medical care is invariably provided.

Community-based accessible population The alternative, community-based accessible populations are composed of individuals in their own homes and defined by a geographic region and a time frame.[4] Samples randomly drawn from community-based accessible populations are unaffected by factors that determine where and when individuals seek health care for a particular problem. Community-based (also called population-based) samples are representative of a geographic region. Because this type of sample is representative of a geographic region, community-based samples are much more accurate than clinic-based samples in determining the need for clinical services in a defined community.

There are, however, a number of problems associ-

ated with community-based samples. Because there are political, economic, and cultural differences among various regions of the country, findings from one region may not necessarily be generalizable to another. Another difficulty with using community-based accessible populations is the time and cost required to generate a sufficiently large number of individuals with a particular disease. The size of the sample required to produce the desired number of subjects with a particular disease depends on the prevalence of the disease in the population. A problem, such as low back pain, that has a high frequency of occurrence in the general population is more practical to study using a community-based sample than is a relatively uncommon disease, such as multiple myeloma. The cost of using a community-based sample is decreased if the required data can be collected by mail or telephone. Because the time and money required to collect data by mail or phone is much lower than that required to collect data in person, it is more economically feasible to draw a large sample when this type of data collection is employed.

In interpreting and applying the findings of a research study, the clinician must judge how well the accessible population represents the target population. For example, a research study with a target population of individuals with occupational low back pain might have used an accessible population of all workers seen in an occupational health clinic at a steel mill in Michigan during 1970–1975. A clinician treating occupational low back pain in migrant workers in southern California in 2001 may decide that the findings from this study are relevant to his practice to the extent that the study dealt with issues that are unaffected by time period (1970s vs 2001) and geography (Michigan vs Southern California). However, it is possible that some of the study findings would not be relevant because there are too many clinically important differences between the 2 time periods and the 2 geographic locations.

SAMPLING STRATEGIES

Once the target and accessible populations are defined by specifying the inclusion and exclusion criteria, the researcher must decide how to select a subset of individuals from the accessible population for inclusion in the study. There are 2 broad categories of sampling techniques, probability sampling and nonprobability sampling (Table 2).[2,4,7]

In probability sampling every member of the accessible population has a specific chance, or probability, of being included in the sample.[4,6,7] Probability samples are created through a process of random selection and, for this reason, are also called random samples.[4,6,7] The other method used to draw samples is nonprobability sampling. In this type of sampling, subjects are chosen on the basis of something other than random selection.[4,7]

TABLE 2. Sampling strategies.

Probability sampling	Nonprobability sampling
Simple random sampling	Convenience sampling
Systematic sampling	Purposive sampling
Stratified random sampling	Snowball sampling
Disproportional sampling	Consecutive sampling
Cluster (multistage) sampling	

Probability Sampling

Although error is always possible when sampling by chance, a randomly drawn sample is considered unbiased in that it has a good chance of reflecting the characteristics of the accessible population from which it was drawn.[1,4,6,7] The size of the sample is more important in reducing the possibility of random sampling error than is the proportion of the population represented by the sample.[6] All of statistical inference, both estimation and hypothesis testing, is based on the assumption that the sample being examined was formed through some type of random process.[1] There are many strategies for producing a random sample. However, all of these strategies require the availability of a list of all the members of the accessible population, which is called a sampling frame.[6] Compiling a valid sampling frame may often be the most difficult and time-consuming aspect of random sampling.[6] Consequently, researchers often use existing sampling frames because they are readily accessible, even if they are less valid. For example, a researcher who wanted to investigate the attitudes and beliefs about evidence-based practice held by physical therapists in the United States could use the readily available list of members of the American Physical Therapy Association (APTA). In doing so, however, the researcher would be making an assumption that the members of the APTA were sufficiently representative of all licensed physical therapists that results, based on the sample of APTA members, could be used to draw inferences concerning all licensed physical therapists. It is up to the clinician to decide whether such assumptions are warranted.

Simple random sampling The most basic type of probability sampling is simple random sampling. In a simple random sample, all subjects have the same chance of participating in the study.[1,4,6,7] This approach requires a random process, such as drawing slips of paper or using a random number table, to draw a sample of a particular size from a valid sampling frame.[4,6,7] For example, if a researcher wanted to investigate the attitudes and beliefs held by physical therapists in the United States about evidence-based practice, the researcher could obtain a list of all licensed physical therapists in the United States by

contacting the licensing body in each state. The researcher would then need to eliminate duplicate entries for individuals licensed in more than 1 state so that all therapists were listed only once, ensuring an equal probability of selection. The final sampling frame would consist of a numbered list of all licensed therapists in the United States. A random number table could then be used to generate a list of numbers of the subjects selected to participate. A list of the therapists linked to the numbers selected would be compiled. The resulting list of therapists would represent the sample for the study.

Systematic sampling Systematic sampling can also be used to draw a random sample from a sampling frame. A systematic sample is formed by first dividing the total number of subjects in the sampling frame by the number of subjects to be selected in the sample. This number represents the interval between selected subjects. The researcher then picks a random starting point and selects periodically consecutive subjects based on the predetermined interval.[4,6,7]

Stratified random and disproportional sampling Sometimes a researcher thinks it is important for certain subject characteristics to be represented in the sample in the same proportion as they are in the accessible population. For example, the researcher might want to ensure that the sample reflects the geographic distribution of physical therapists in the United States. If 10% of physical therapists in the United States reside in California, then 10% of the physical therapists in the sample should reside in California. Proportionate representation can be accomplished using stratified random sampling.[4,6,7] In the case of stratified random sampling, the sampling frame is divided into strata, or groups, based on some characteristic. In our example, the characteristic of interest would be state of residence. In using this technique, the researcher would first determine the percentage of physical therapists residing in each state. Given the percentage for each state and the overall sample size, the researcher would then calculate the number of therapists that should be selected from each state. The researcher would then use either simple random sampling or systematic sampling to select the correct number of subjects from each state. The final sample would be a random sample that was guaranteed to have the same proportion of subjects from each state as the sampling frame (accessible population).

A similar approach can be used when the researcher is concerned that subjects with certain characteristics will not appear in the sample in sufficient number to allow for adequate comparisons. If subgroups within the population are considerably unequal in size, random sampling may produce too few subjects within certain subgroups. For example, a researcher might want to compare the attitudes and beliefs of physical therapists living in more urban states to those who live in more rural states. However, she might be concerned that a simple random sample would contain too few subjects from rural states because those states contribute a relatively smaller number of therapists to the total number of licensed physical therapists in the United States. In order to ensure that the sample contained an adequate number of physical therapists from rural states, the researcher could employ a technique known as disproportional sampling. In disproportional sampling, the sampling frame is divided into subgroups based on the characteristic of interest.[4,6,7] In this example, the researcher would have to decide which states were urban and which were rural and group them accordingly. The researcher then would draw a sample of equal size from each group. Because one subgroup (rural) is much smaller than the other (urban), this sample will not reflect the true population proportion of the therapists in these subgroups. Therefore, if a sample is formed in this manner, the attitudes and beliefs of the rural group will be overrepresented in the sample when it is examined as a whole. The effect of disproportional sampling can be controlled by weighting the data so that the contribution of members of various subgroups reflects their proportion in the sampling frame (accessible population). For example, if the mean age of therapists in the rural subgroup were much younger than the mean age of therapists in the urban group, an attempt to estimate the average age of physical therapists licensed in the United States using the type of disproportional sample described would produce an estimate that was younger than it should be because rural therapists contributed more to the sample average than to the population average. This problem can be corrected by using the appropriate proportional weights. The weights are calculated by determining the probability that any 1 member of a subgroup would be selected if simple random sampling were used. The weight assigned to a subject in a particular subgroup is the inverse of the probability that a member of that subgroup would be selected using simple random sampling. When these weights are used, the sample means are unbiased estimates of the population means.

Cluster sampling or multistage sampling Sometimes, it is simply impossible to construct a sampling frame at the individual subject level. For example, it would be very difficult for a researcher to investigate the attitudes and beliefs about evidence-based practice held by physical therapists in the United States practicing in the acute care setting because there is no available sampling frame. However, it is still possible to select a random sample of these therapists using cluster sampling, or multistage sampling.[4,6,7] This form of sampling depends on linking members of the target population to some established groupings that can be sampled. For example, it is possible to

compile a list of all the counties in the United States and then use a random process to select a specific number of counties. Once the counties were selected, a list of the acute care hospitals in each county could be generated using public information. The researcher could then compile a list of acute care hospitals across all the sampled counties and select a random sample of hospitals from this composite list. Then using the random sample of hospitals, the researcher could contact the director of rehabilitation at each hospital to obtain a list of all physical therapists practicing in acute care. Assuming the therapists working at the hospitals agreed to have their names released, a list of all physical therapists practicing in acute care at the sampled hospitals could be compiled and a random sample of therapists could be selected from the list. Therapists selected would constitute a random sample of physical therapists practicing in acute care settings in the United States.

Nonprobability Sampling

Convenience sampling A convenience, or accidental sample, is formed by taking those members of the accessible population who are easily available.[4,7] For example, consider an accessible population consisting of patients attending a university-based osteoarthritis clinic Monday and Wednesday mornings from 7:00 to 11:00 AM and Friday afternoons from 1:00 to 5:00 PM. A researcher who has time available on Friday afternoons for recruiting subjects may enroll only subjects who attend the clinic on Friday afternoons. The subjects he recruits would comprise a convenience sample. A problem with convenience samples is that they are biased relative to the very factors that made the subjects convenient. Subjects who schedule clinic visits on Friday afternoon might differ on factors such as age, employment status, disease severity, and even religion from subjects who schedule early Monday or Wednesday morning visits. A sample drawn in this manner would not necessarily represent all of the patients seen by this clinic.

Volunteers represent another kind of convenience sample. Studies that recruit subjects by advertising for volunteers may form samples that differ substantially from the accessible population from which these volunteers are drawn. Volunteers may tend to be both mentally and physically healthier than nonvolunteers and may also tend to have fewer work and care giving responsibilities.[3] The nature of the bias introduced by using volunteers in a study largely depends on the research question. It is important to note that the subjects enrolled in clinical trials are always volunteers,[3] and there may be many important differences between individuals who will agree to participate in a clinical trial and those who will not. Although the use of volunteers in a clinical trial may sometimes limit generalization of the findings, random assignment to treatment or control condition is a much more important determinant of the validity of a clinical trial than is random selection of subjects.[3,7]

Purposive sampling Purposive, or judgmental, sampling represents another type of nonprobability sampling. This type of sampling involves purposely selecting those members of the accessible population judged to be most appropriate for the study, even though the sample selected may not be representative of either the accessible or target populations.[4,7] For example, subjects may be selected because they are known to be very tolerant, compliant, or cooperative. Unfortunately, findings that an intervention is effective or well-tolerated by this type of subject may not be very useful in deciding whether to use the intervention with a more usual group of patients.

Snowball sampling Snowball sampling is another form of nonprobability sampling. Subjects recruited during the early phases of the study are asked to help identify other individuals who might meet the criteria for admission to the study.[7] Because this type of sampling depends on social contacts, it tends to produce a group of subjects who are fairly similar to one another and, therefore, less representative of the entire accessible population.

Consecutive sampling A consecutive sample involves enrolling every subject who meets the selection criteria during a specified time interval.[4,7] Use of consecutive sampling is essentially the same as enrolling the entire accessible population. Such a sample should be highly representative of the accessible population, unless there are periodic fluctuations in the type of patient seen at the clinical site and the time frame for recruitment is too short to accommodate these fluctuations. If a clinical research study utilizes consecutive sampling, then the clinician only needs to decide how well the accessible population represents the target population in order to judge the applicability of the research findings.

SAMPLING BIAS

A probability sample is considered unbiased in that it has a good chance of reflecting the characteristics of the accessible population from which it was drawn.[1,4,6,7] In contrast, nonprobability samples are frequently subject to sampling bias. Sampling bias occurs when the individuals selected for a sample either over- or underrepresent population attributes that are related to the phenomenon of interest.[4,6,7] For example, if purposive sampling is used to select subjects who will strictly adhere to a home exercise program and if strict adherence to the program is critical to produce the desired change in functional mobility, then the sample will be biased in a way that will affect the outcome of the study. If, in fact, the typical patient does not follow the home

exercise program as closely as the research subjects did, then the typical patients may achieve less satisfactory results than did the study participants because the subjects used in the study do not well represent the type patient for whom the program may be prescribed.

Even probability samples can have problems with sampling bias if too many subjects withdraw from the sample. Sampling bias also occurs when there is a high nonresponse or drop out rate in a study.[4,6,7] Although the researchers may have used random sampling to draw the initial sample, if more than approximately 25% of the sample drops out or fails to respond to a phone or mail questionnaire, the remaining portion of the sample may be biased. Frequently, subjects who drop out or fail to respond differ in important ways from subjects who do not. Removing subjects from the sample who have characteristics associated with not responding or dropping out causes those characteristics to be underrepresented in the remaining sample. The sample is, therefore, no longer representative of the accessible population. For example, if a researcher was studying natural history of occupational low back pain by following a cohort of workers for 5 years, a certain proportion of the workers would probably be lost to follow-up before the study was completed. If a large proportion of the subjects lost to follow-up disappeared because they stopped working due to back pain, the findings of the study would be severely biased.

SUMMARY

How does a clinician decide whether to use the findings of a research study to assist in clinical decision making for an individual patient? It is necessary to determine if the patient seeking care is a member of the target population for the study. This requires examining whether the patient meets the inclusion and exclusion criteria specified by the researchers. If the patient is a member of the target population, then the clinician must decide if it is appropriate to draw conclusions about the target population based on the findings of the study. It is important to determine if the researchers adhered strictly to their own inclusion and exclusion criteria. It is also important to determine if a large proportion of the subjects dropped out or failed to respond, leaving a biased sample.

If it appears that the sample used in the study is acceptably close to the planned sample, the clinician must decide whether the sample represents the population of interest fairly. It is important to determine if the researchers used some form of probability sampling and if they did not, what biases were introduced? It is also important to determine whether there were differences between the accessible population and the target population that might be relevant to the individual patient?

If the clinician decides that the findings of the study may be appropriately generalized to the target population and if the clinician's patient is a member of that target population, then the clinician can use the findings of the research study to assist in making clinical decisions concerning this individual patient.

REFERENCES

1. Daniel WW. Biostatistics: A Foundation for Analysis in The Health Sciences. New York, NY: John Wiley & Sons; 1999.
2. Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology: The Essentials.* Baltimore, Md: Williams & Wilkins; 1996.
3. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials.* 2nd ed. St. Louis, Mo: Mosby Year Book; 1985.
4. Hulley SB, Cummings SR. *Designing Clinical Research.* Baltimore, Md: Williams & Wilkins; 1988.
5. Kelsey JL, Thompson WD, Evans AS. Methods of sampling and estimation of sample size. In: *Methods in Observational Epidemiology.* New York, NY: Oxford University Press; 1986: 254–284.
6. Levy PS, Lemeshow S. *Sampling for Health Professionals.* Belmont, Calif: Lifetime Learning Publications; 1980.
7. Portney LG, Watkins MP. Sampling. In: *Foundations of Clinical Research: Application to Practice.* Upper Saddle River, NJ: Prentice Hall Health; 2000.
8. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* Boston, Md: Little Brown; 1991.
9. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based Medicine: How to Practice & Teach EBM.* New York, NY: Churchill Livingstone; 1997.