SEVENTH EDITION

THE PRACTICE OF NURSING RESEARCH

Appraisal, Synthesis, and Generation of Evidence

SUSAN K. **GROVE**
NANCY **BURNS**
JENNIFER R. **GRAY**

ELSEVIER

http://evolve.elsevier.com

CHAPTER 15

Sampling

Many of us have preconceived notions about samples and sampling, which we acquired from television commercials, polls of public opinion, market researchers, and newspaper reports of research findings. The advertiser boasts that four of five doctors recommend its product; the newscaster announces that John Jones is predicted to win the senate election by a margin of 3 to 1; the newspaper reports that scientists' studies have found that taking a statin drug, such as atorvastatin (Lipitor), significantly reduces the risk of coronary artery disease.

All of these examples use sampling techniques. However, some of the outcomes are more valid than others, partly because of the sampling techniques used. In most instances, television, newspapers, and advertisements do not explain their sampling techniques. You may hold opinions about the adequacy of these techniques, but there is not enough information to make a judgment.

The sampling component is an important part of the research process that needs to be carefully thought out and clearly described. To achieve these goals, researchers need to understand the techniques of sampling and the reasoning behind them. With this knowledge, you can make intelligent judgments about sampling when you are critically appraising studies or developing a sampling plan for your own study. This chapter examines sampling theory and concepts; sampling plans; probability and nonprobability sampling methods for quantitative, qualitative, outcomes, and intervention research; sample size; and settings for conducting studies. The chapter concludes with a discussion of the process for recruiting and retaining subjects or participants for study samples in various settings.

Sampling Theory

Sampling involves selecting a group of people, events, behaviors, or other elements with which to conduct a study. A **sampling plan** defines the process of making the sample selections; **sample** denotes the selected group of people or elements included in a study. Sampling decisions have a major impact on the meaning and generalizability of the findings.

Sampling theory was developed to determine mathematically the most effective way to acquire a sample that would accurately reflect the population under study. The theoretical, mathematical rationale for decisions related to sampling emerged from survey research, although the techniques were first applied to experimental research by agricultural scientists. One of the most important surveys that stimulated improvements in sampling techniques was the U.S. census. Researchers have adopted the assumptions of sampling theory identified for the census surveys and incorporated them within the research process (Thompson, 2002).

Key concepts of sampling theory are (1) populations, (2) elements, (3) sampling criteria, (4) representativeness, (5) sampling errors, (6) randomization, (7) sampling frames, and (8) sampling plans. The following sections explain these concepts; later in the chapter, these concepts are used to explain various sampling methods.

Populations and Elements

The **population** is a particular group of people, such as people who have had a myocardial infarction, or type of element, such as nasogastric tubes, that is the focus of the research. The **target population** is the entire set of individuals or elements who meet the sampling criteria, such as women who have experienced a myocardial infarction in the past year. Figure 15-1 shows the relationships among the population, target population, and accessible populations. An **accessible population** is the portion of the target population to which the researchers have reasonable access. The accessible population might be elements

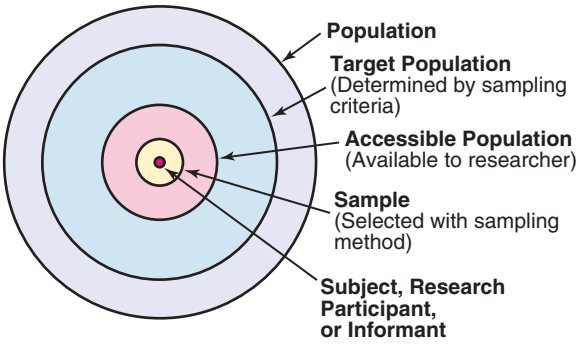

Figure 15-1 Population, sample, and subject selected for a study.

within a country, state, city, hospital, nursing unit, or clinic, such as the adults with diabetes in a primary care clinic in Fort Worth, Texas. The sample is obtained from the accessible population by a particular sampling method, such as simple random sampling. The individual units of the population and sample are called **elements**. An element can be a person, event, behavior, or any other single unit of study. When elements are persons, they are usually referred to as **subjects** or **research participants** or **informants** (see Figure 15-1). The term used by researchers depends of the philosophical paradigm that is reflected in the study and the design. The term *subject,* and sometimes *research participant,* is used within the context of the postpositivist paradigm of quantitative research (see Chapter 2). The term *study* or *research participant* or *informant* is used in the context of the naturalistic paradigm of qualitative research (Fawcett & Garity, 2009; Munhall, 2012). In quantitative, intervention, and outcomes research, the findings from a study are generalized first to the accessible population and then, if appropriate, more abstractly to the target population.

Generalizing means that the findings can be applied to more than just the sample under study because the sample is representative of the target population. Because of the importance of generalizing, there are risks to defining the accessible population too narrowly. For example, a narrow definition of the accessible population reduces the ability to generalize from the study sample to the target population and diminishes the meaningfulness of the findings. Biases may be introduced that make generalization to the broader target population difficult to defend. If the accessible population is defined as individuals in a white, upper-middle-class setting, one cannot generalize to nonwhite or lower income populations. These biases are similar to biases that may be encountered in a nonrandom sample (Thompson, 2002).

In some studies, the entire population is the target of the study. These studies are referred to as **population studies** (Barhyte, Redman, & Neill, 1990). Many of these studies use data available in large databases, such as the census data or other government-maintained databases. Epidemiologists sometimes use entire populations for their large database studies. In other studies, the entire population of interest in the study is small and well defined. For example, one could conduct a study in which the defined population was all living recipients of heart and lung transplants.

In some cases, a hypothetical population is defined for a study. A **hypothetical population** assumes the presence of a population that cannot be defined according to sampling theory rules, which require a list of all members of the population. For example, individuals who successfully lose weight would be a hypothetical population. The number of individuals in the population, who they are, how much weight they have lost, how long they have kept the weight off, and how they achieved the weight loss are unknown. Some populations are elusive and constantly changing. For example, identifying all women in active labor in the United States, all people grieving the loss of a loved one, or all people coming into an emergency department would be impossible.

Sampling or Eligibility Criteria

Sampling criteria, also referred to as **eligibility criteria**, include a list of characteristics essential for membership or eligibility in the target population. The criteria are developed from the research problem, the purpose, a review of literature, the conceptual and operational definitions of the study variables, and the design. The sampling criteria determine the target population, and the sample is selected from the accessible population within the target population (see Figure 15-1). When the study is complete, the findings are generalized from the sample to the accessible population and then to the target population if the study has a representative sample (see the next section).

You might identify broad sampling criteria for a study, such as all adults older than 18 years of age able to read and write English. These criteria ensure a large target population of **heterogeneous** or diverse potential subjects. A heterogeneous sample increases your ability to generalize the findings to a larger target population. In descriptive or correlational studies, the sampling criteria may be defined to ensure a heterogeneous population with a broad range of values for

the variables being studied. However, in quasi-experimental or experimental studies, the primary purpose of sampling criteria is to limit the effect of extraneous variables on the particular interaction between the independent and dependent variables. In these types of studies, the sampling criteria need to be specific and designed to make the population as **homogeneous** or similar as possible to control for the extraneous variables. Subjects are selected to maximize the effects of the independent variable and minimize the effects of variation in other extraneous variables so that they have a limited impact on the dependent variable scores.

Sampling criteria may include characteristics such as the ability to read, to write responses on the data collection instruments or forms, and to comprehend and communicate using the English language. Age limitations are often specified, such as adults 18 years and older. Subjects may be limited to individuals who are not participating in any other study. Persons who are able to participate fully in the procedure for obtaining informed consent are often selected as subjects. If potential subjects have diminished autonomy or are unable to give informed consent, consent must be obtained from their legal representatives. Thus, persons who are legally or mentally incompetent, terminally ill, or confined to an institution are more difficult to access as subjects (see Chapter 9). However, sampling criteria should not become so restrictive that the researcher cannot find an adequate number of study participants.

A study might have inclusion or exclusion sampling criteria (or both). **Inclusion sampling criteria** are characteristics that a subject or element must possess to be part of the target population. **Exclusion sampling criteria** are characteristics that can cause a person or element to be excluded from the target population. Researchers need to provide logical reasons for their inclusion and exclusion sampling criteria, and certain groups should not be excluded without justification. In the past, some groups, such as women, ethnic minorities, elderly adults, and poor people, were unnecessarily excluded from studies (Larson, 1994). Today, federal funding for research is strongly linked to including these populations in studies. Exclusion criteria limit the generalization of the study findings and should be carefully considered before being used in a study.

Twiss et al. (2009) conducted a quasi-experimental study to examine the effects of strength and weight training (ST) exercises on muscle strength, balance, and falls of breast cancer survivors (BCSs) with bone loss (population). This study included clearly identified inclusion and exclusion sampling or eligibility criteria that are presented in the following excerpt.

> "Women were included if they were 35-77 years of age, had a history of stage 0 (in situ), I, or II breast cancer, a BMD [bone mineral density] T-score of −1.0 or less at any of three sites (hip, spine, forearm), were at least 6 months post breast-cancer treatment and 12 months postmenopausal, resided within 100 miles of one of four research sites (Omaha, Lincoln, Kearney, and Scottsbluff, NE), and had their physicians' permission to participate [inclusion sampling criteria]. Women were excluded if they (a) had a recurrence of breast cancer; (b) were currently taking hormone therapy, bisphosphonates, glucocorticosteroids, or other drugs affecting bone; (c) were currently engaging in ST exercises; (d) had a body mass index (BMI) of 35 or greater; (e) had serum calcium, creatinine, or thyroid stimulating hormone (if on thyroid therapy) outside normal limits; or (f) had active gastrointestinal problems or other conditions that prohibited ST exercises, risedronate, calcium, or vitamin D intake [exclusion sampling criteria]." (Twiss et al., 2009, p. 72)

Twiss et al. (2009) identified specific inclusion and exclusion sampling criteria to designate the subjects in the target population precisely. These sampling criteria probably were narrowly defined by the researchers to promote the selection of a homogeneous sample of postmenopausal BCSs with bone loss. These inclusion and exclusion sampling criteria were appropriate for the study to reduce the effect of possible extraneous variables that might have an impact on the treatment (ST exercises) and the measurement of the dependent variables (muscle strength, balance, and falls). Because this is a quasi-experimental study that examined the impact of the treatment on the dependent or outcome variables, the increased controls imposed by the sampling criteria strengthened the likelihood that the study outcomes were caused by the treatment and not by extraneous variables. Twiss et al. (2009) found significant improvement in muscle strength and balance for the treatment group but no significant difference in the number of falls between the treatment and comparison groups.

Sample Representativeness

For a sample to be **representative**, it must be similar to the target population in as many ways as possible. It is especially important that the sample be

representative in relation to the variables you are studying and to other factors that may influence the study variables. For example, if your study examines attitudes toward acquired immunodeficiency syndrome (AIDS), the sample should represent the distribution of attitudes toward AIDS that exists in the specified population. In addition, a sample must represent the demographic characteristics, such as age, gender, ethnicity, income, and education, which often influence study variables.

The accessible population must be representative of the target population. If the accessible population is limited to a particular setting or type of setting, the individuals seeking care at that setting may be different from the individuals who would seek care for the same problem in other settings or from individuals who self-manage their problems. Studies conducted in private hospitals usually exclude poor patients, and other settings could exclude elderly or undereducated patients. People who do not have access to care are usually excluded from health-focused studies. Subjects and the care they receive in research centers are different from patients and the care they receive in community clinics, public hospitals, veterans' hospitals, and rural health clinics. Obese individuals who choose to enter a program to lose weight may differ from obese individuals who do not enter a program. All of these factors limit representativeness and limit our understanding of the phenomena important in practice.

Representativeness is usually evaluated by comparing the numerical values of the sample (a **statistic** such as the mean) with the same values from the target population. A numerical value of a population is called a **parameter**. We can estimate the **population parameter** by identifying the values obtained in previous studies examining the same variables. The accuracy with which the population parameters have been estimated within a study is referred to as **precision**. Precision in estimating parameters requires well-developed methods of measurement that are used repeatedly in several studies. You can define parameters by conducting a series of descriptive and correlational studies, each of which examines a different segment of the target population; then perform a meta-analysis to estimate the population parameter (Thompson, 2002).

Sampling Error

The difference between a sample statistic and a population parameter is called the **sampling error** (Figure 15-2). A large sampling error means that the sample is not providing a precise picture of the population; it is not representative. Sampling error is usually larger

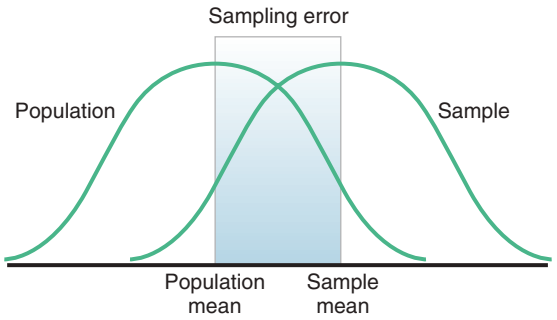

Figure 15-2 Sampling error.

with small samples and decreases as the sample size increases. Sampling error reduces the **power** of a study, or the ability of the statistical analyses conducted to detect differences between groups or to describe the relationships among variables (Aberson, 2010; Cohen, 1988). Sampling error occurs as a result of random variation and systematic variation.

Random Variation

Random variation is the expected difference in values that occurs when one examines different subjects from the same sample. If the mean is used to describe the sample, the values of individuals in that sample will not all be exactly the same as the sample mean. Values of individual subjects vary from the value of the sample mean. The difference is random because the value of each subject is likely to vary in a different direction. Some values are higher and others are lower than the sample mean. The values are randomly scattered around the mean. As the sample size becomes larger, overall variation in sample values decreases, with more values being close to the sample mean. As the sample size increases, the sample mean is also more likely to have a value similar to that of the population mean.

Systematic Variation

Systematic variation, or **systematic bias**, is a consequence of selecting subjects whose measurement values are different, or vary, in some specific way from the population. Because the subjects have something in common, their values tend to be similar to the values of others in the sample but different in some way from the values of the population as a whole. These values do not vary randomly around the population mean. Most of the variation from the mean is in the same direction; it is systematic. All the values in the sample may tend to be higher or lower than the mean of the population (Thompson, 2002).

For example, if all the subjects in a study examining some type of healthcare knowledge have an intelligence quotient (IQ) higher than 120, many of their scores will likely be higher than the mean of a population that includes individuals with a wide variation in IQ, such as IQs that range from 90 to 130. The IQs of the subjects have introduced a systematic bias. This situation could occur, for example, if all the subjects were college students, which has been the case in the development of many measurement methods in psychology.

Because of systematic variance, the sample mean is different from the population mean. The extent of the difference is the sampling error (see Figure 15-2). Exclusion criteria tend to increase the systematic bias in the sample and increase the sampling error. An extreme example of this problem is the highly restrictive sampling criteria used in some experimental studies that result in a large sampling error and greatly diminished representativeness.

If the method of selecting subjects produces a sample with a systematic bias, increasing the sample size would not decrease the sampling error. When a systematic bias occurs in an experimental study, it can lead the researcher to believe that a treatment has made a difference when, in actuality, the values would be different even without the treatment. This situation usually occurs because of an interaction of the systematic bias with the treatment.

Refusal and Acceptance Rates in Studies

Systematic variation or bias is most likely to occur when the sampling process is not random. However, even in a random sample, systematic variation can occur if potential subjects decline participation. Systematic bias increases as the subjects' refusal rate increases. A **refusal rate** is the number and percentage of subjects who declined to participate in the study. High refusal rates to participate in a study have been linked to individuals with serious physical and emotional illnesses, low socioeconomic status, and weak social networks (Neumark, Stommel, Given, & Given, 2001). The higher the refusal rate, the less the sample is representative of the target population. The refusal rate is calculated by dividing the number of potential subjects refusing to participate by the number of potential subjects meeting sampling criteria and multiplying the results by 100%.

Refusal rate formula = number potential subjects refusing to participate ÷ number potential subjects meeting sample criteria × 100%

For example, if 200 potential subjects met the sampling criteria, and 40 refused to participate in the study, the refusal rate would be 20%.

Refusal rate = 40 (number refusing) ÷ 200 (number meeting sampling criteria) = 0.2 × 100% = 20%

Sometimes researchers provide an **acceptance rate**, or the number and percentage of the subjects who agree to participate in a study, rather than a refusal rate. The acceptance rate is calculated by dividing the number of potential subjects who agree to participate in a study by the number of potential subjects who meet sampling criteria and multiplying the result by 100%.

Acceptance rate formula = number potential subjects agreeing to participate ÷ number potential subjects meeting sample criteria × 100%

If you know the refusal rate, you can also subtract the refusal rate from 100% to obtain the acceptance rate. Usually researchers report either the acceptance rate or the refusal rate but not both. In the example mentioned earlier, 200 potential subjects met the sampling criteria; 160 agreed to participate in the study, and 40 refused.

Acceptance rate = 160 (number accepting) ÷ 200 (number meeting sampling criteria) = 0.8 × 100% = 80%

Acceptance rate = 100% − refusal rate or 100% − 20% = 80%

Sample Attrition and Retention Rates in Studies

Systematic variation can also occur in studies with high sample attrition. **Sample attrition** is the withdrawal or loss of subjects from a study. Systematic variation is greatest when a high number of subjects withdraw from the study before the data have been collected or when a large number of subjects withdraw from one group but not the other in the study (Kerlinger & Lee, 2000; Thompson, 2002). In studies involving a treatment, subjects in the control group who do not receive the treatment may be more likely to withdraw from the study. Sample attrition should be reported in the published study to determine if the final sample represents the target population. Researchers also need to provide a rationale for subjects withdrawing from the study and to determine if they are

different from the subjects who complete the study. The sample is most like the target population if the attrition rate is low (<10% to 20%) and the subjects withdrawing from the study are similar to the subjects completing the study. Sample attrition rate is calculated by dividing the number of subjects withdrawing from a study by the **sample size** and multiplying the results by 100%.

Sample attrition rate formula = number subjects withdrawing ÷ sample size × 100%

For example, if a study had a sample size of 160, and 40 people withdrew from the study, the attrition rate would be 25%.

Attrition rate = 40 (number withdrawing) ÷ 160 (sample size) = 0.25 × 100% = 25%

The opposite of the attrition rate is the **retention rate**, or the number and percentage of subjects completing the study. The higher the retention rate, the more representative the sample is of the target population, and the more likely the study results are an accurate reflection of reality. Often researchers identify either the attrition rate or the retention rate but not both. It is better to provide a rate in addition to the number of subjects withdrawing or completing a study. In the example just presented with a sample size of 160, if 40 subjects withdrew from the study, then 120 subjects were retained or completed the study. The retention rate is calculated by dividing the number of subjects completing the study by the initial sample size and multiplying by 100%.

Sample retention rate formula = number subjects completing study ÷ sample size × 100%

Retention rate = 120 (number retained) ÷ 160 (sample size) = 0.75 × 100% = 75%

The study by Twiss et al. (2009) of the effects of ST exercises on muscle strength, balance, and falls of BCSs with bone loss was introduced earlier in this chapter with the discussion of sampling criteria; the following excerpt presents the acceptance rate and sample attrition for this study.

> "A sample of 249 participants met the screening criteria and they were enrolled in the study.… Of the 249 women, 223 completed the 24-month testing and were included in the analysis (exercise [treatment group] = 110; comparison = 113). The remaining 26 women (exercise = 14; comparison = 12) withdrew from the study before 24 months. Reasons for withdrawal included the desire for a different exercise program (n = 7); insufficient time (n = 6); intolerance to meds (n = 5); cancer recurrence (n = 5); health problems (n = 2); and relocation (n = 1)." (Twiss et al., 2009, p. 22)

Twiss et al. (2009) identified that 249 participants or subjects met the sampling criteria and 249 were enrolled in the study indicating that the acceptance rate for the study was 100%. The sample retention was 223 women for a retention rate of 90% (223 ÷ 249 × 100% = 89.6% = 90%), and the sample attrition rate was 26 women for an attrition rate of 10% (100% − 90% = 10%). The treatment group retention was 110 women with a retention rate of 89% (110 ÷ 124 × 100% = 88.7% = 89%). The comparison group retention was 113 women with a retention rate of 90% (113 ÷ 125 = 90.4% = 90%). This study has an excellent acceptance rate (100%) and a very strong sample retention rate of 90% for a 24-month-long study. The retention rates for both groups were very strong and comparable (treatment group 89% and comparison group 90%). Twiss et al. (2009) also provided a rationale for the subjects' attrition, and the reasons were varied and seemed appropriate and typical for a study lasting 24 months. The acceptance rate, the sample and group retention rates, and the reasons for subjects' attrition indicate limited potential for systematic variation in the study sample. The likelihood is increased that the sample is representative of the target population and the results are an accurate reflection of reality. The study would have been strengthened if the researchers would have included not only the numbers but also the sample and group retention rates.

Randomization

From a sampling theory point of view, **randomization** means that each individual in the population should have a greater than zero opportunity to be selected for the sample. The method of achieving this opportunity is referred to as **random sampling**. In experimental studies that use a control group, subjects are randomly selected and randomly assigned to either the control group or the experimental group. The use of the term **control group**—the group not receiving the treatment—is usually limited to studies using random sampling and random assignment to the treatment and control groups. The control group usually receives no care. If nonrandom sampling methods are used for sample selection, the group not receiving a treatment

receives usual or standard care and is generally referred to as a **comparison group**. With a comparison group, there is an increase in the possibility of preexisting differences between that group and the experimental group receiving the treatment.

Random sampling increases the extent to which the sample is representative of the target population. However, random sampling must take place in an accessible population that is representative of the target population. Exclusion criteria limit true randomness. Thus, a study that uses random sampling techniques may have such restrictive sampling criteria that the sample is not truly random. In any case, it is rarely possible to obtain a purely random sample for nursing studies because of informed consent requirements. Even if the original sample is random, persons who volunteer or consent to participate in a study may differ in important ways from persons who are unwilling to participate. All samples with human subjects must be **volunteer samples**, which includes individuals willing to participate in the study, to protect the rights of the individuals (Fawcett & Garity, 2009). Methods of achieving random sampling are described later in the chapter.

Sampling Frame

For each person in the target or accessible population to have an opportunity to be selected for the sample, each person in the population must be identified. To accomplish this goal, the researcher must acquire a list of every member of the population through the use of the sampling criteria to define membership. This listing of members of the population is referred to as the **sampling frame**. The researcher selects subjects from the sampling frame using a sampling plan. Djukic, Kovner, Budin, and Norman (2010) studied the effect of nurses' perceived physical work environment on their job satisfaction and described their sampling frame in the following excerpt.

> "The study was conducted at a large urban hospital in the U.S. northeast region that is a nongovernment, not-for-profit, general medical and surgical major teaching hospital. About 1,300 staff RNs [population] were employed at the hospital at the time of the study.… A total of 746 RNs who met eligibility criteria were invited to participate in the study [sampling frame of target population]. The eligible RNs were those who had a functioning work e-mail account and who worked fulltime, on inpatient units, providing direct patient care." (Djukic et al., 2010, pp. 444-445)

The sampling frame in this study included the names of the 746 RNs who were asked to participate in the study.

Sampling Plan

A **sampling plan** describes the strategies that will be used to obtain a sample for a study. The plan is developed to enhance representativeness, reduce systematic bias, and decrease the sampling error. Sampling strategies have been devised to accomplish these three tasks and to optimize sample selection. The sampling plan may use probability (random) sampling methods or nonprobability (nonrandom) sampling methods.

A **sampling method** is the process of selecting a group of people, events, behaviors, or other elements that represent the population being studied. A sampling method is similar to a design; it is not specific to a study. The sampling plan provides detail about the application of a sampling method in a specific study. The sampling plan must be described in detail for purposes of critical appraisal, replication, and future meta-analyses. The sampling method implemented in a study varies with the type of research being conducted. Quantitative, outcomes, and intervention research apply a variety of probability and nonprobability sampling methods. Qualitative research usually includes nonprobability sampling methods. The sampling methods to be included in this text are identified in Table 15-1 and are linked to the types of research that most commonly incorporate them. The following sections describe the different types of probability and nonprobability sampling methods most commonly used in quantitative, qualitative, outcomes, and intervention research in nursing.

Probability (Random) Sampling Methods

Probability sampling methods have been developed to ensure some degree of precision in estimations of the population parameters. Probability samples reduce sampling error. The term **probability sampling method** refers to the fact that every member (element) of the population has a probability higher than zero of being selected for the sample. Inferential statistical analyses are based on the assumption that the sample from which data were derived has been obtained randomly. Thus, probability sampling methods are often referred to as **random sampling methods**. These samples are more likely to represent the population than samples obtained with nonprobability sampling methods. All subsets of the population, which may differ from one another but contribute to the

TABLE 15-1 Probability and Nonprobability Sampling Methods Commonly Applied in Nursing Research

Sampling Method	Common Applications
Probability Sampling Methods	
Simple random sampling	Quantitative, outcomes, and intervention research
Stratified random sampling	Quantitative, outcomes, and intervention research
Cluster sampling	Quantitative, outcomes, and intervention research
Systematic sampling	Quantitative, outcomes, and intervention research
Nonprobability Sampling Methods	
Convenience sampling	Quantitative, qualitative, outcomes, and intervention research
Quota sampling	Quantitative, outcomes, and intervention research
Purpose or purposeful sampling	Qualitative and sometimes quantitative research
Network or snowball sampling	Qualitative and sometimes quantitative research
Theoretical sampling	Qualitative research

parameters of the population, have a chance to be represented in the sample. Probability sampling methods are most commonly applied in quantitative, outcomes, and intervention research.

There is less opportunity for systematic bias if subjects are selected randomly, although it is possible for a systematic bias to occur by chance. Using random sampling, the researcher cannot decide that person *X* would be a better subject for the study than person *Y*. In addition, a researcher cannot exclude a subset of people from selection as subjects because he or she does not agree with them, does not like them, or finds them hard to deal with. Potential subjects cannot be excluded just because they are too sick, not sick enough, coping too well, or not coping adequately. The researcher, who has a vested interest in the study, could (consciously or unconsciously) select subjects whose conditions or behaviors are consistent with the study hypothesis. It is tempting to exclude uncooperative or assertive individuals. Random sampling leaves the selection to chance and decreases sampling error and increases the validity of the study (Thompson, 2002).

Theoretically, to obtain a probability sample, the researcher must develop a sampling frame that includes every element in the population. The sample must be randomly selected from the sampling frame. According to sampling theory, it is impossible to select a sample randomly from a population that cannot be clearly defined. Four sampling designs have been developed to achieve probability sampling: simple random sampling, stratified random sampling, cluster sampling, and systematic sampling.

Simple Random Sampling

Simple random sampling is the most basic of the probability sampling methods. To achieve simple random sampling, elements are selected at random from the sampling frame. This goal can be accomplished in various ways, limited only by the imagination of the researcher. If the sampling frame is small, the researcher can write names on slips of paper, place the names in a container, mix well, and draw out one at a time until the desired sample size has been reached. Another technique is to assign a number to each name in the sampling frame. In large population sets, elements may already have assigned numbers. For example, numbers are assigned to medical records, organizational memberships, and professional licenses. The researcher can use a computer to select these numbers randomly to obtain a sample.

There can be some differences in the probability for the selection of each element, depending on whether the name or number of the selected element is replaced before the next name or number is selected. Selection with replacement, the most conservative random sampling approach, provides exactly equal opportunities for each element to be selected (Thompson, 2002). For example, if the researcher draws names out of a hat to obtain a sample, each name must be replaced before the next name is drawn to ensure equal opportunity for each subject.

Selection without replacement gives each element different levels of probability for selection. For example, if the researcher is selecting 10 subjects from a population of 50, the first name has a 1 in 5 chance (10 draws, 50 names), or a 0.2 probability, of being selected. If the first name is not replaced, the remaining 49 names have a 9 in 49 chance, or a 0.18 probability, of being selected. As further names are drawn, the probability of being selected decreases.

There are many ways to achieve random selection, such as with the use of a computer, a random numbers table, drawing names out of a hat, or a roulette wheel. The most common method of random selection is the computer, which can be programmed to select a sample randomly from the sampling frame with replacement. However, some researchers still use a table of random numbers to select a random sample.

TABLE 15-2 Section from a Random Numbers Table

06	84	10	22	56	72	25	70	69	43
07	63	10	34	66	39	54	02	33	85
03	19	63	93	72	52	13	30	44	40
77	32	69	58	25	15	55	38	19	62
20	01	94	54	66	88	43	91	34	28

Table 15-2 shows a section from a random numbers table. To use a table of random numbers, the researcher places a pencil or a finger on the table with the eyes closed. The number touched is the starting place. Moving the pencil or finger up, down, right, or left, the researcher uses the numbers in order until the desired sample size is obtained. For example, the researcher places a pencil on 58 in Table 15-2, which is in the fourth column from the left and fourth row down. If five subjects are to be selected from a population of 100 and the researcher decides to go across the column to the right, the subject numbers chosen are 58, 25, 15, 55, and 38. Table 15-2 is useful only if the population number is less than 100. However, tables are available for larger populations, such as the random numbers table provided in the online resources for this textbook or the Thompson (2002, pp. 14-15) sampling text.

Degirmen, Ozerdogan, Sayiner, Kosgeroglu, and Ayranci (2010, p. 153) conducted a pretest-posttest randomized controlled experimental study to determine the effect of hand and foot massage and foot massage only interventions on the postoperative pain of women who had a cesarean operation. These researchers obtained their sample using a simple random sampling method that is described in the following excerpt from their study.

> "The study was conducted in obstetric intensive care units and services of all the public and university hospitals in the province of Eskisehir, Turkey.... During the 4 month study, 281 patients attended for the cesarean operations to the obstetric intensive care units and services of all hospitals concerned [target population and settings]. The total 75 study patients [sample] out of the 281 were selected by random sampling method from the patients' presenting orders [sampling frame] and evenly divided into three groups; a control group, a foot and hand massage group, and a foot massage group, each of which included 25 patients.... Because some patients accepted the intervention before the operation, but changed their mind after the operation (3 patients in total), not all patients participated in the study." (Degirmen et al., 2010, p. 154)

Degirmen et al. (2010) clearly identified their target population as women needing cesarean operations, and the 281 women with presenting orders provided the sampling frame for the study. The sample of 75 women was randomly selected, but the researchers did not indicate the process for the random selection. The use of a computer to select a sample randomly is usually the most efficient and unbiased process. The subjects were evenly divided with 25 in each group, but the researchers do not indicate if the assignment to groups was random or based on the convenience of the subjects or researchers. Application of simple random sampling and the attrition of only three (4%) subjects from the study seem to provide a sample representative of the target population. However, the study would have been strengthened by a discussion of the process for random sampling and a clarification of how the subjects were assigned to groups. The outcomes of the study were that foot and hand massage interventions significantly reduced postoperative pain experienced by the women and that foot and hand massage was significantly more effective than foot massage only.

Stratified Random Sampling

Stratified random sampling is used when the researcher knows some of the variables in the population that are critical to achieving representativeness. Variables commonly used for stratification are age, gender, ethnicity, socioeconomic status, diagnosis, geographical region, type of institution, type of care, care provider, and site of care. The variable or variables chosen for stratification need to be correlated with the dependent variables being examined in the study. Subjects within each stratum are expected to be more similar (homogeneous) in relation to the study variables than they are to be similar to subjects in other strata or the total sample. In stratified random sampling, the subjects are randomly selected on the basis of their classification into the selected strata.

For example, if in conducting your research you selected a stratified random sample of 100 adult subjects using age as the variable for stratification, the sample might include 25 subjects in the age range 18 to 39 years, 25 subjects in the age range 40 to 59 years, 25 subjects in the age range 60 to 79 years, and 25

subjects 80 years or older. Stratification ensures that all levels of the identified variable, in this example age, are adequately represented in the sample. With a stratified random sample, you could use a smaller sample size to achieve the same degree of representativeness as a large sample acquired through simple random sampling. Sampling error decreases, power increases, data collection time is reduced, and the cost of the study is lower if stratification is used (Fawcett & Garity, 2009; Thompson, 2002).

One question that arises in relation to stratification is whether each stratum should have equivalent numbers of subjects in the sample (termed **disproportionate sampling**) or whether the numbers of subjects should be selected in proportion to their occurrence in the population (termed **proportionate sampling**). For example, if stratification is being achieved by ethnicity and the population is 45% white non-Hispanic, 25% Hispanic nonwhite, 25% African American, and 5% Asian, your research team would have to decide whether to select equal numbers of each ethnic group or to calculate a proportion of the sample. Good arguments exist for both approaches. Stratification is not as useful if one stratum contains only a small number of subjects. In the aforementioned situation, if proportions are used and the sample size is 100, the study would include only five Asians, hardly enough to be representative. If equal numbers of each group are used, each group would contain at least 25 subjects; however, the white non-Hispanic group would be underrepresented. In this case, mathematically weighting the findings from each stratum can equalize the representation to ensure proportional contributions of each stratum to the total score of the sample. Most textbooks on sampling describe this procedure (Levy & Lemsbow, 1980; Thompson, 2002; Yates, 1981).

Ulrich et al. (2006) used a stratified random sampling method to obtain their sample of nurse practitioners (NPs) and physician assistants (PAs) for the purpose of studying the ethical conflict of these healthcare providers associated with managed care. The following excerpt from this study describes the sampling method used to obtain the final sample of 1536 providers (833 NPs and 689 PAs).

> "A self-administered questionnaire was mailed to an initial stratified random sample [sampling method] of 3,900 NPs and PAs practicing in the United States. The sample was selected from the national lists provided by Medical Marketing Services, an independently owned organization that manages medical industry lists (www.mmslists.com/main.asp). The list for PAs was derived from the American Academy of Physicians Assistants (AAPA), and a comprehensive list of NPs was derived from the medical and nursing boards of the 50 states and the District of Columbia [sampling frames for NPs and PAs].… After undeliverable (1.9%) and other disqualified respondents (13.2%, i.e., no longer practicing, non-primary-care practitioner) were removed, the overall adjusted response rate was 50.6%." (Ulrich et al., 2006, p. 393)

The study sampling frames for the NPs and PAs are representative of all 50 states and the District of Columbia, and the lists for the sampling frames were from quality sources. The study has a strong response rate of 50.6% for a mailed questionnaire, and the researchers identified why certain respondents were disqualified. The final sample was large (1536 subjects) with strong representation for both NPs (833 subjects) and PAs (689 subjects). The study sample might have been stronger with a more equal number of NP and PA subjects. The 833 NPs and 689 PAs add to 1522 subjects and it is unclear why the sample size is identified as 1536 unless there are missing data from subjects. However, the sample was a great strength of this study and appeared to represent the target population of NPs and PAs currently practicing in primary care in the United States.

Cluster Sampling

Cluster sampling is a probability sampling method applied when the population is heterogeneous; it is similar to stratified random sampling but takes advantage of the natural clusters or groups of population units that have similar characteristics (Fawcett & Garity, 2009). Cluster sampling is used in two situations. The first situation is when a simple random sample would be prohibitive in terms of travel time and cost. Imagine trying to arrange personal meetings with 100 people, each in a different part of the United States. The second situation is in cases in which the individual elements making up the population are unknown, preventing the development of a sampling frame. For example, there is no list of all the heart surgery patients who complete rehabilitation programs in the United States. In these cases, it is often possible to obtain lists of institutions or organizations with which the elements of interest are associated.

In cluster sampling, the researcher develops a sampling frame that includes a list of all the states, cities,

institutions, or organizations with which elements of the identified population would be linked. States, cities, institutions, or organizations are selected randomly as units from which to obtain elements for the sample. In some cases, this random selection continues through several stages and is referred to as **multistage cluster sampling**. For example, the researcher might first randomly select states and next randomly select cities within the sampled states. Hospitals within the randomly selected cities might then be randomly selected. Within the hospitals, nursing units might be randomly selected. At this level, either all the patients on the nursing unit who fit the criteria for the study might be included, or patients could be randomly selected.

Cluster sampling provides a means for obtaining a larger sample at a lower cost. However, it has some disadvantages. Data from subjects associated with the same institution are likely to be correlated and not completely independent. This correlation can cause a decrease in precision and an increase in sampling error. However, such disadvantages can be offset to some extent by the use of a larger sample.

Fouladbakhsh and Stommel (2010, p. E8) used multistage cluster sampling in their study of the "complex relationships among gender, physical and psychological symptoms, and use of specific CAM [complementary and alternative medicine] health practices among individuals living in the United States who have been diagnosed with cancer." These researchers described their sampling method in the following excerpt from their study.

> "The NHIS [National Health Interview Survey] methodology employs a multistage probability cluster sampling design [sampling method] that is representative of the NHIS target universe, defined as 'the civilian noninstitutionalized population' (Botman, Moore, Moriarty, & Parsons, 2000, p. 14; National Center for Health Statistics). In the first stage, 339 primary sampling units were selected from about 1,900 area sampling units representing counties, groups of adjacent counties, or metropolitan areas covering the 50 states and the District of Columbia [1st stage cluster sampling]. The selection included all of the most populous primary sampling units in the United States and stratified probability samples (by state, area poverty level, and population size) of the less populous ones. In a second step, primary sampling units were partitioned into substrata (up to 21) based on concentrations of African American and Hispanic populations [2nd stage cluster sampling]. In a third step, clusters of dwelling units form the secondary sampling units selected from each substratum [3rd stage cluster sampling]. Finally, within each secondary sampling unit, all African American and Hispanic households were selected for interviews, whereas other households were sampled at differing rates within the substrata. Therefore, the sampling design of the NHIS includes oversampling of minorities." (Fouladbakhsh & Stommel, 2010, pp. E8-E9)

These researchers detailed their use of multistage cluster sampling and clearly identified the three stages implemented and the rationale for each stage. The study had a large, national sample that seemed representative of all 50 states and the District of Columbia with an oversampling of minorities to accomplish the purpose of the study. The complex cluster sampling method used in this study provided a representative sample, which decreases the likelihood of sampling error and increases the validity of the study findings. The findings reported by Fouladbakhsh and Stommel (2010, p. E7) indicated that "CAM practice use was more prevalent among female, middle-aged, Caucasian, and well-educated subjects. Pain, depression, and insomnia were strong predictors of practice use, with differences noted by gender and practice type."

Systematic Sampling

Systematic sampling can be conducted when an ordered list of all members of the population is available. The process involves selecting every *k*th individual on the list, using a starting point selected randomly. If the initial starting point is not random, the sample is not a probability sample. To use this design in your research, you must know the number of elements in the population and the size of the sample desired. Divide the population size by the desired sample size, giving *k,* the size of the gap between elements selected from the list. For example, if the population size is $N = 1200$ and the desired sample size is $n = 100$, then you could calculate the value of k:

$$k = \text{population size} \div \text{sample size desired}$$

Example: $k = 1200$ (population size) $\div$ 100 (sample size desired) $= 12$

Thus, $k = 12$, which means that every 12th person on the list would be included in the sample. Some authors

argue that this procedure does not truly give each element an opportunity to be included in the sample; it provides a random but unequal chance for inclusion (Thompson, 2002).

Researchers must be careful to determine that the original list has not been set up with any ordering that could be meaningful in relation to the study. The process is based on the assumption that the order of the list is random in relation to the variables being studied. If the order of the list is related to the study, systematic bias is introduced. In addition to this risk, it is difficult to compute sampling error with the use of this design (Floyd, 1993).

Li and Mukamel (2010, p. S256) conducted a secondary data analysis of National Nursing Home Survey (NNHS) data to identify any "racial disparities in the receipt and documentation of influenza and pneumococcus vaccinations among nursing-home residents." The NNHS data were obtained using stratified random sampling and systematic sampling for selection of the nursing homes and random sampling of the residents for interviews. The researchers described the sampling plan for the NNHS and the participants for their study in the following study excerpt.

> "We obtained the public-use resident file of the Centers for Disease Control and Prevention's 2004 NNHS, which includes a nationally representative sample of nursing homes, their residents, and the services the residents received.… The 2004 NNHS involved a stratified 2-stage probability design. The first stage was the selection of nursing homes stratified by geographical location (state, county, and zip code), bed-size category (number of beds), and ownership status (profit vs. nonprofit). Nursing homes were finally selected by systematic sampling, with the probability proportional to bed-size category. The second stage, sampling of current residents, was carried out by interviewers at the time of their visits to the facilities. Individuals were randomly selected from patient rosters, and a sample of up to 12 current residents per facility was selected for the final interview. The final NNHS comprised 13,507 residents in 1174 nursing homes, with an overall response rate of 78%.… Our analyses were limited to 11,448 nonHispanic Whites and 1384 Blacks, excluding 765 residents (5.7%) of other races/ethnicities.… The final sample comprised 10,562 residents for analysis of influenza and 12,134 residents for the analysis of pneumococcus vaccinations." (Li & Mukamel, 2010, p. S256)

Li and Mukamel (2010) used a national database developed with strong probability sampling methods (stratified random sampling and systematic sampling). The database included a nationally representative sample of nursing homes and their residents, which decreases the potential for sampling error and supports the validity of the findings. The study design was limited to an examination of data from white and black subjects, who were the largest groups represented in the nursing homes surveyed. The sample size was extremely large, which increases the potential to generalize the findings to these two ethnic groups. The study would have been strengthened by a more detailed discussion of the systematic sampling of the nursing homes and the random sampling of the residents. The researchers found that disparities existed in vaccination coverage of white and black nursing home residents and recommended interventions to improve coverage.

Nonprobability (Nonrandom) Sampling Methods Commonly Applied in Quantitative Research

In **nonprobability sampling**, not every element of the population has an opportunity to be included in the sample. Nonprobability sampling methods increase the likelihood of obtaining samples that are not representative of their target populations. However, most nursing studies use nonprobability sampling, especially convenience sampling, to select study samples. In conducting studies in nursing and other health disciplines, limited subjects are available, and it is often impossible to obtain a random sample. Researchers often include any subjects willing to participate who meet the eligibility criteria.

There are several types of nonprobability (nonrandom) sampling designs. Each addresses a different research need. The five nonprobability sampling designs described in this textbook are (1) convenience sampling, (2) quota sampling, (3) purposive or purposeful sampling, (4) network or snowball sampling, and (5) theoretical sampling. These sampling methods are applied in both quantitative and qualitative research. However, convenience sampling and quota sampling are applied more often in quantitative, outcomes, and intervention research than in qualitative studies and are discussed in this section (see Table 15-1). Purposive sampling, network sampling, and theoretical sampling are more commonly applied in

qualitative studies than in quantitative studies and are discussed later in this chapter.

Convenience Sampling

In **convenience sampling**, subjects are included in the study because they happened to be in the right place at the right time. Researchers simply enter available subjects into the study until they have reached the desired sample size. Convenience sampling, also called **accidental sampling**, is considered a weak approach to sampling because it provides little opportunity to control for biases. Multiple biases may exist in convenience sampling; these biases range from minimal to serious. Researchers need to identify and describe known biases in their samples. You can identify biases by carefully thinking through the sample criteria used to determine the target population and taking steps to improve the representativeness of the sample. For example, in a study of home care management of patients with complex healthcare needs, educational level would be an important extraneous variable. One solution for controlling this extraneous variable would be to redefine the sampling criteria to include only patients with a high school education. Doing so would limit the extent of generalization but decrease the bias created by educational level. Another option would be to select a population known to include individuals with a wide variety of educational levels. Data could be collected on educational level so that the description of the sample would include information on educational level. With this information, one could judge the extent to which the sample was representative with respect to educational level (Thompson, 2002).

Decisions related to sample selection must be carefully described to enable others to evaluate the possibility of biases. In addition, data need be gathered to allow a thorough description of the sample that can also be used to evaluate for possible biases. Data on the sample can be used to compare the sample with other samples and to estimate the parameters of populations through meta-analyses.

Many strategies are available for selecting a convenience sample. A classroom of students might be used. Patients who attend a clinic on a specific day, subjects who attend a support group, patients currently admitted to a hospital with a specific diagnosis, and every fifth person who enters the emergency department are examples of types of commonly selected convenience samples.

Convenience samples are inexpensive and accessible, and they usually require less time to acquire than other types of samples. Convenience samples provide means to conduct studies on topics that could not be examined through the use of probability sampling. Convenience sampling enables researchers to acquire information in unexplored areas. According to Kerlinger and Lee (2000), a convenience sample is probably not that bad when it is used with reasonable knowledge and care in implementing a study. Healthcare studies are usually conducted with particular types of patients experiencing varying numbers of health problems; these patients often are reluctant to participate in research. Thus, researchers often find it very difficult to recruit subjects for their studies and frequently must use a sample of convenience versus random sampling to obtain their sample.

O'Shea, Wallace, Griffin, and Fitzpatrick (2011, p. 35) used convenience sampling to determine the "effectiveness of a spiritual educational session on pediatric nurses' perspectives concerning the provision of pediatric spiritual care." The following excerpt describes their sampling method.

> "The setting for the data collection was a large university-affiliated children's hospital located in the northeast. Participants represented a convenience sample [sampling method] from a potential 355 registered neonatal and pediatric staff nurses employed at the hospital [target population]. Forty-one nurses voluntarily consented to participate. Number of participants per session varied from approximately two to five, depending on the day and time of the session." (O'Shea et al., 2011, p. 37)

O'Shea et al. (2011) clearly identified their sampling method, target population, and sample size. The acceptance rate for the study was 41 neonatal and pediatric nurses, which is only 12% of the 355 nurses in the target population. The nurses volunteering to participate in the study might be different in some way from the nurses refusing to participate. The small sample size ($N = 41$) and low acceptance rate increase the chance for sampling error and decrease the representativeness of the sample. However, all 41 participants completed this quasi-experimental study (0% attrition rate), which decreases the potential for bias. The sample was homogeneous for registered nurse status, employed on pediatric units, and gender (female $n = 40$). The sample was heterogeneous for education ranging from associate's degree to master's in nursing and for years of experience ranging from less than 2 years to more than 20 years. In a quasi-experimental study, a homogeneous sample decreases the extraneous variables that might influence the findings. Based on the sampling method (nonprobability convenience

sample), small sample size, high refusal rate, and differences in the education and years of experience of the nurses, the findings from this study are best generalized to the sample and not the accessible or target populations. The researchers found that the educational sessions had a positive effect on the nurses' perspectives toward providing spiritual care, but additional research is needed to confirm the effect of this intervention. Additional studies with large convenience samples that have similar results would indicate the effectiveness of this intervention for practice.

Quota Sampling

Quota sampling uses a convenience sampling technique with an added feature, a strategy to ensure the inclusion of subject types or strata in a population that are likely to be underrepresented in the convenience sample, such as women, minority groups, elderly adults, poor people, rich people, and undereducated adults. This method may also be used to mimic the known characteristics of the target population or to ensure adequate numbers of subjects in each stratum for the planned statistical analyses. The technique is similar to the technique used in stratified random sampling, but the initial sample is not random. If necessary, mathematical weighting can be used to adjust sample values so that they are consistent with the proportion of subgroups found in the population. Quota sampling offers an improvement over convenience sampling and tends to decrease potential biases. In most studies in which convenience samples are used, quota sampling could be used and should be considered (Thompson, 2002).

Pieper, Templin, Kirsner, and Birk (2010) used quota sampling to examine the impact of vascular leg disorders, such as chronic venous disorders and peripheral arterial disease, on the physical activity levels of opioid-addicted adults in a methadone-maintained program. The following excerpt describes their sampling process.

> "The sample ($N = 713$) was obtained from September 2005 to December 2007 from 12 methadone treatment clinics located in a large urban area [convenience sampling]. The sample was stratified on four variables: age (25-39 years, 40-49 years, 50-65 years); gender (male, female); ethnicity (African American, White); and drug use (nonIDU [injection drug use], arm/upper body injection only, or legs ± upper body injection; Pieper, Templin, Kirsner, & Birk, 2009) [quota sampling]. The purpose of the stratification was to allow comparisons of type of drug use with minimal confounding by age, gender, or ethnicity. Additional inclusion criteria included presence of both legs, able to walk, and able to speak and understand English. The analyses reported here are on the 569 participants who completed the revised LDUQ [Legs in Daily Use Questionnaire], which were edited after examining the test-retest data from 104 participants, not included in the 569, who were tested first." (Pieper et al., 2010, p. 429)

Pieper et al. (2010) clearly identified that the original sample was one of convenience because it included people attending 12 methadone treatment clinics who were willing to participate in the study. The quota sampling involved stratification of the sample on four variables with a clear rationale for the variables selected for stratification. The study was completed by 569 participants, but 104 participants were used for examining the test-retest reliability of the LDUQ and not included in the final data collection. The study had an attrition of 40 participants (attrition rate = $40 \div 569 \times 100\% = 7\%$).

The study by Pieper et al. (2010) has several strengths in the sampling process. The use of quota sampling ensured that the study sample was more representative of the target population than using convenience sampling only. In addition, the participants were obtained from 12 different clinics and the sample size was large ($N = 713 - 104$ [used only for instrument reliability testing] = 569) with a small attrition rate (7%). The sample appeared to be representative of the target population with limited potential for sampling error. The findings indicated that motivation was the strongest predictor of physical activity and that healthcare professionals need to evaluate the vascular health of legs of drug injection users before encouraging exercise.

Nonprobability Sampling Methods Commonly Applied in Qualitative Research

Qualitative research is conducted to gain insights and discover meaning about a particular experience, situation, cultural element, or historical event. The intent is an in-depth understanding of a purposefully selected sample and not the generalization of the findings from a randomly selected sample to a target population, as in quantitative, outcomes, and intervention research. In

qualitative research, experiences, events, and incidents are more the focus of sampling than people (Marshall & Rossman, 2011; Munhall, 2012; Patton, 2002). The researcher attempts to select participants or informants who can provide extensive information about the experience or event being studied. For example, if the goal of your study was to describe the phenomenon of living with chronic pain, you would purposefully select participants who were articulate and reflective, had a history of chronic pain, and were willing to share their chronic pain experience (Coyne, 1997).

The three common sampling methods applied in qualitative research are purposive or purposeful sampling, network or snowball sampling, and theoretical sampling (see Table 15-1). These sampling methods enable the researcher to select the specific participants who would provide the most extensive information about the phenomenon, event, or situation being studied (Marshall & Rossman, 2011). The sample selection process can have a profound effect on the quality of the research and should be described in enough depth to promote the interpretation of the findings and the replication of the study (Munhall, 2012; Patton, 2002).

Purposive Sampling

In **purposive sampling**, sometimes referred to as *purposeful, judgmental,* or *selective sampling,* the researcher consciously selects certain participants, elements, events, or incidents to include in the study. In purposive sampling, qualitative researchers select **information-rich cases**, or cases that can teach them a great deal about the central focus or purpose of the study (Green & Thorogood, 2004; Patton, 2002). Efforts might be made to include typical and atypical participants or situations. Researchers also seek **critical cases**, or cases that make a point clearly or are extremely important in understanding the purpose of the study (Munhall, 2012). The researcher might select participants or informants of various ages, participants with differing diagnoses or illness severity, or participants who received an ineffective treatment versus an effective treatment for their illness.

This sampling plan has been criticized because it is difficult to evaluate the precision of the researcher's judgment. How does one determine that the patient or element was typical or atypical, good or bad, effective or ineffective? Researchers need to indicate the characteristics that they desire in participants and provide a rationale for selecting these types of participants to obtain essential data for their study. Purposive sampling method is used in qualitative research to gain insight into a new area of study or to obtain in-depth understanding of a complex experience or event (Munhall, 2012).

Sternberg and Barry (2011, p. 64) applied purposive and snowball sampling methods in conducting their phenomenological study of "the experiences of transnational Latina mothers who immigrated to the United States without legal documentation or their children." Snowball sampling, discussed in the next section, involves current study participants identifying additional potential study participants who are similar to them. Sternberg and Barry describe their sampling methods in the following study excerpt.

> "A purposive sample of Latina mothers who immigrated to the United States without their children was selected. Interviews were conducted and analyzed until saturation and redundancy was achieved (Munhall, 2007). Eight women in total participated in the study. Three women visiting a free community clinic in southeast Florida were invited to participate in the study by the nurse researcher; the other five were recruited using snowball sampling. Snowball sampling was chosen over random sampling because of the difficulty in gaining access to a population with so many undocumented members (Munhall, 2007). Further, it enabled the researcher to establish a trusting relationship with the participants and obtain a more heterogeneous sample group. To be included in this study, participants had to be mothers who were 18 years of age or older, Spanish or English speaking, and immigrants to the United States from Latin America who had left their child or children in their country of origin." (Sternberg & Barry, 2011, pp. 65-66)

Sternberg and Barry (2011) clearly identified their sampling methods that were appropriate for the qualitative study they conducted. The initial three participants were identified through purposive sampling so that they could achieve a group reflective of their sampling criteria. Individuals without legal documentation are hard to locate, so use of snowball sampling was appropriate to identify five additional participants. The eight participants provided an adequate-sized sample because the researchers were able to reach saturation and redundancy of themes during their data analysis. The findings from the study included seven essential themes: "living in extreme poverty, having hope, choosing to walk from poverty, suffering through the trip to and across the U.S.-Mexican border, mothering from afar, valuing family,

and changing personally" (Sternberg & Barry, 2011, p. 67).

Network (Snowball) Sampling

Network sampling, sometimes referred to as *snowball* or *chain sampling,* holds promise for locating samples difficult or impossible to obtain in other ways or that had not been previously identified for study. Network sampling takes advantage of social networks and the fact that friends tend to have characteristics in common. When you have found a few participants with the necessary criteria, you can ask for their assistance in getting in touch with others with similar characteristics. The first few participants are often obtained through convenience or purposive sampling methods, and the sample size is expanded using network or snowball sampling. This sampling method is occasionally used in quantitative studies, but it is more commonly used in qualitative studies. In qualitative research, network sampling is an effective strategy for identifying participants who know other potential participants who can provide the greatest insight and essential information about an experience or event that is identified for study (Marshall & Rossman, 2011; Munhall, 2012; Patton, 2002).

This strategy is also particularly useful for finding participants in socially devalued populations, such as alcoholics, child abusers, sex offenders, drug addicts, and criminals. These individuals are seldom willing to make themselves known. Other groups, such as widows, grieving siblings, or individuals successful at lifestyle changes, can be located using this strategy. These individuals are outside the existing healthcare system and are difficult to find. Biases are built into the sampling process because the participants are not independent of one another. However, the participants selected have the expertise to provide the essential information needed to address the study purpose. The study by Sternberg and Barry (2011) presented in the previous section applied snowball or network sampling to identify additional study participants who had immigrated to the United States without legal documentation. These researchers clearly identified their use of snowball sampling and their rationale for using this method in their study.

Theoretical Sampling

Theoretical sampling is usually applied in grounded theory research to advance the development of a selected theory throughout the research process (Munhall, 2012). The researcher gathers data from any individual or group that can provide relevant data for theory generation. The data are considered relevant if they include information that generates, delimits, and saturates the theoretical codes in the study needed for theory generation. A code is saturated if it is complete and the researcher can see how it fits in the theory. The researcher continues to seek sources and gather data until the codes are saturated and the theory evolves from the codes and the data. Diversity in the sample is encouraged so that the theory developed covers a wide range of behavior in varied situations (Marshall & Rossman, 2011; Patton, 2002).

Beaulieu, Kools, Kennedy, and Humphreys (2011, p. 41) conducted a qualitative study using grounded theory methods to "explore and better understand the reasons for the apparent underuse of emergency contraceptive pills (ECPs) in young people in coupled relationships." These researchers applied three sampling methods: (1) convenience sampling, (2) snowball sampling, and (3) theoretical sampling. They described their sampling methods in the following study excerpt.

> "A convenience sample was recruited via public notices and snowball sampling (Fain, 2004). Inclusion criteria were women 18 to 25 years of age, English speaking, with basic knowledge of ECPs, and currently involved in a sexual relationship with a partner who was also willing to participate in the study.… Analysis began simultaneously with data collection as dictated by the tenets of grounded theory.… The initial analysis of interviews and field notes consisted of strategies of open coding and memoing (Glaser & Strauss, 1967).… As new categories emerged, the original interview guide was revised and additional couples were recruited to allow for theoretical sampling, that is, sampling specifically to fill in theoretical gaps, strengthen categories and their relationships, and verify or challenge emerging conceptualizations (Strauss & Corbin, 1998) and forced coding (Charmaz, 2006)." (Beaulieu et al., 2011, p. 43)

Beaulieu et al. (2011) clearly identified their sampling methods that were appropriate for a qualitative study conducted with grounded theory methodology. Both convenience and snowball sampling methods were applied because the researchers wanted an adequate number of couples to participate in their study and discuss their decision making regarding ECPs use. Beaulieu et al. also provided a detailed rationale for their use of theoretical sampling to develop a theory about decision making of young couples related to ECPs use. The sampling methods provided a strong sample of 22 couples, who provided the

essential information for grounded theory development (Glaser & Strauss, 1967; Strauss & Corbin, 1998). More details on this study are presented later in this chapter in the discussion of sample size in qualitative studies.

Sample Size in Quantitative Research

One of the questions beginning researchers commonly ask is, "What size sample should I use?" Historically, the response to this question has been that a sample should contain at least 30 subjects for each study variable measured. Statisticians consider 30 subjects as the minimum number for data on a single variable to approach a normal distribution. So if a study includes 4 variables, researchers need at least 120 subjects in their final sample. Researchers are encouraged to determine the possible attrition rate for their study to ensure an adequate sample size at the completion of their study. For example, researchers might anticipate a 10-15% attrition rate in their study and need to obtain a sample of 130 to 140 subjects to ensure the final sample size after attrition is 120. The best method of determining sample size is a power analysis but if information is not available to conduct a power analysis, this recommendation of 30 subjects per study variable might be used.

The deciding factor in determining an adequate sample size for correlational, quasi-experimental, and experimental studies is power. **Power** is the capacity of the study to detect differences or relationships that actually exist in the population. Expressed another way, power is the capacity to reject a null hypothesis correctly. The minimum acceptable power for a study is commonly recommended to be 0.80 (80%) (Aberson, 2010; Cohen, 1988; Kraemer & Thiemann, 1987). If you do not have sufficient power to detect differences or relationships that exist in the population, you might question the advisability of conducting the study. You determine the sample size needed to obtain sufficient power by performing a **power analysis**. Power analysis includes the standard power (usually 80%), level of significance (usually set at 0.05 in nursing studies), effect size (discussed in the next section), and sample size.

An increasing number of nurse researchers are using power analysis to determine sample size, but it is essential that the results of the power analyses be included in the published studies. Not conducting a power analysis for a study and omitting the power analysis results in a published study are significant problems if the study failed to detect significant differences or relationships, which might be due to an inadequate sample size. The calculation of the power analysis varies with the types of statistical analyses used to analyze study data. Statistical programs are available to conduct a power analysis for a study (see Chapter 21). However, you can get a general idea about sample size using the power tables in Appendix F in this textbook.

The adequacy of sample sizes must be evaluated more carefully in future nursing studies before data collection. Studies with inadequate sample sizes should not be approved for data collection unless they are preliminary pilot studies conducted before a planned larger study. If it is impossible for you to obtain a larger sample because of time or numbers of available subjects, you should redesign your study so that the available sample is adequate for the planned analyses. If you cannot obtain a sufficient sample size, you should not conduct the study.

Large sample sizes are difficult to obtain in nursing studies, require long data collection periods, and are costly. In developing the methodology for a study, you must evaluate the elements of the methodology that affect the required sample size. Kraemer and Thiemann (1987) identified the following factors that must be taken into consideration in determining sample size:

1. The more stringent the significance level (e.g., 0.001 versus 0.05), the greater the necessary sample size. Most nursing studies include a level of significance or alpha (α) = 0.05.
2. Two-tailed statistical tests require larger sample sizes than one-tailed tests. (Tailedness of statistical tests is explained in Chapters 21 and 25.)
3. The smaller the effect size, the larger the necessary sample size. The effect size is a determination of the effectiveness of a treatment on the outcome (dependent) variable or the strength of the relationship between two variables.
4. The larger the power required, the larger the necessary sample size. Thus, a study requiring a power of 90% requires a much larger sample than a study with power set at 80%.
5. The smaller the sample size, the smaller the power of the study (Aberson, 2010; Cohen, 1988; Kraemer & Thiemann, 1987).

The factors that must be considered in decisions about sample size (because they affect power) are effect size, type of study, number of variables, sensitivity of the measurement methods, and data analysis techniques. These factors are discussed in the following sections.

Effect Size

Effect is the presence of a phenomenon. If a phenomenon exists, it is not absent, and the null hypothesis is in error. However, effect is best understood when not considered in a dichotomous way—that is, as either present or absent. If a phenomenon exists, it exists to some degree. **Effect size** (*ES*) is the extent to which a phenomenon is present in a population. In this case, the term *effect* is used in a broader sense than the term *cause and effect*. For example, you might examine the impact of distraction on the experience of pain during an injection. To examine this question, you might obtain a sample of subjects receiving injections and measure the perception of pain in a group of subjects who were distracted during injection and a group of subjects who were not distracted. The null hypothesis would be: "There is no difference in the level of pain perceived by the treatment group receiving distraction than the comparison group receiving no distraction." If this were so, you would say that the effect of distraction on the perception of pain was zero, and the null hypothesis would be accepted. In another study, you might be interested in using the Pearson product moment correlation *r* to examine the relationship between coping and anxiety. Your null hypothesis is that the population *r* would be zero, or coping is not related to anxiety (Cohen, 1988).

In a study, it is easier to detect large differences between groups than to detect small differences. Strong relationships between variables in a study are easier to detect than weak relationships. Thus, smaller samples can detect large *ESs*; smaller *ESs* require larger samples. *ESs* can be positive or negative because variables are positively and negatively correlated. A negative *ES* is calculated if a treatment causes a decrease in the study mean, such as an exercise program that decreases the weight of subjects. Broadly speaking, the definitions for *ES* strengths might be as follows:

Small *ES* would be <0.3 or <–0.3
Medium *ES* would be about 0.3 to 0.5 or –0.3 to –0.5
Large *ES* would be >0.5 or >–0.5

These broad ranges are provided because the *ES* definitions of small, medium, and large vary based on the analysis being conducted. For example, the *ESs* for comparing two means, such as the treatment group mean with the comparison group mean (expressed as *d*), are small = 0.2 or –0.2, medium = 0.5 or –0.5, and large = 0.8 or –0.8. The *ESs* for relationships (expressed as *r*) might be defined as small = 0.1 or –0.1, medium = 0.3 or –0.3, and large = 0.5 or –0.5 (Aberson, 2010; Cohen, 1988).

Extremely small *ESs* (e.g., <0.1) may not be clinically important because the relationships between the variables are small or the differences between the treatment and comparison groups are limited. Knowing the *ES* that would be regarded as clinically important allows us to limit the sample to the size needed to detect that level of *ES* (Kraemer & Thiemann, 1987). A result is clinically important if the effect is large enough to alter clinical decisions. For example, in comparing glass thermometers with electronic thermometers, an effect size of 0.1° F in oral temperature is probably not important enough to influence selection of a particular type of thermometer in clinical practice. The clinical importance of an *ES* varies on the basis of the variables being studied and the population.

ESs vary according to the population being studied. Researchers must determine the *ES* for the particular relationship or effect being studied in a selected population. The most desirable source of this information is evidence from previous studies (Aberson, 2010; Melnyk & Fineout-Overholt, 2011). The correlation value (*r*) is equal to the *ES* for the relationship between two variables. For example, if depression is correlated with anxiety at $r = 0.45$, then the $ES = r = 0.45$, a medium *ES*.

ES* formula for relationships = *r

Example: $ES = r = 0.45$

In published studies with treatments, means and standard deviations can be used to calculate the *ES* (Grove, 2007). For example, if the mean weight loss for the treatment or intervention group is 5 pounds per month with a standard deviation (*SD*) = 4.5, and the mean weight loss of the control or comparison group is 1 pound per month with *SD* = 6.5, you can calculate the *ES*, which is usually expressed as *d*.

***ES* formula for group differences = *d* = mean of the treatment group – mean of the control group ÷ standard deviation of control group**

Example: $ES = d = 5 - 1 \div 6.5 = 4 \div 6.5 = 0.615 = 0.62$

This calculation can be used only as an estimate of *ES* for the study. If the researcher changes the measurement method used, the design of the study, or the population being studied, the *ES* will be altered. When estimating *ES* based on previous studies, you might note the *ESs* vary from 0.33 to 0.45; it is best to choose the lower *ES* of 0.33 to calculate a sample size for a study. The best estimate of a population parameter of

ES is obtained from a meta-analysis in which an estimated population *ES* is calculated through the use of statistical values from all studies included in the analysis (Aberson, 2010; Cohen, 1988).

If few relevant studies have been conducted in the area of interest, a small pilot study can be performed, and data analysis results can be used to calculate the *ES*. If pilot studies are not feasible, dummy power table analysis can be used to calculate the smallest *ES* with clinical or theoretical value. Yarandi (1991) described the process of calculating a dummy power table. If all else fails, *ES* can be estimated as small, medium, or large. Numerical values would be assigned to these estimates and the power analysis performed. Cohen (1988) and Aberson (2010) indicated the numerical values for small, medium, and large effects on the basis of specific statistical procedures. In new areas of research, *ESs* for studies are usually set as small (<0.3) (Aberson, 2010).

Jones, Duffy, and Flanagan (2011) conducted a randomized clinical trial to test the efficacy of a nurse-coached telephone intervention on the distress symptoms and functional health status of ambulatory arthroscopic surgery patients. These researchers conducted a power analysis to identify the sample size for their study, and it is described in the following excerpt.

> "The inclusion criteria were as follows: adults (18 years or older) who were able to read and write English, would undergo ambulatory arthroscopic surgery under general anesthesia, had telephone access at home, and were discharged home on the day of surgery [target population]. According to power calculations, a sample size of 102 participants would assure a power >.80, given a significance level [alpha] of .05, three measurement times, a minimum correlation of repeated measure of .30, and a low to moderate effect size (d = .75)." (Jones et al., 2011, p. 94)

Jones et al. (2011) clearly identified their target population and the process for determining their sample size using power analysis. The standard power of 0.80 or 80% was used, and alpha was set at 0.05, which is common in nursing studies. The focus of the study was determining differences between the treatment and comparison groups, so the *ES* was expressed as $d = 0.75$, which is a moderate ES for examining differences between groups (see previous discussion of *ESs*). Jones et al. (2011, p. 92) found the "intervention participants had significantly less symptom distress at 72 hours and 1 week postsurgery and significantly better overall physical and mental health at 1 week postsurgery than those who received usual practice." The significant results indicate the study had an adequate sample size to determine differences between the intervention or experimental group and usual practice or comparison group. If the study findings had been nonsignificant, the researchers would have needed to conduct a power analysis to determine the power achieved in the study.

Type of Study

Descriptive case studies tend to use small samples. Groups are not compared, and problems related to sampling error and generalization have little relevance for such studies. A small sample size may better serve the researcher who is interested in examining a situation in depth from various perspectives. Other descriptive studies, particularly studies using survey questionnaires, and correlational studies often require large samples. In these studies, multiple variables may be examined, and extraneous variables are likely to affect subject responses to the variables under study. Statistical comparisons are often made among multiple subgroups in the sample, requiring that an adequate sample be available for each subgroup being analyzed. In addition, subjects are likely to be heterogeneous in terms of demographic variables, and measurement tools are sometimes not adequately refined. Although target populations may have been identified, sampling frames may be unavailable, and parameters have not usually been well defined by previous studies. All of these factors decrease the power of the study and require increases in sample size (Aberson, 2010; Kraemer & Thiemann, 1987).

In the past, quasi-experimental and experimental studies often used smaller samples than descriptive and correlational studies. As control in the study increases, the sample size can decrease and still approximate the population. Instruments in these studies tend to be refined, improving precision. However, sample size must be sufficient to achieve an acceptable level of power (0.8) and reduce the risk of a type II error (indicating the study findings are nonsignificant, when they are really significant) (Aberson, 2010; Kraemer & Thiemann, 1987).

The study design influences power, but the design with the greatest power may not always be the most valid design to use. The experimental design with the greatest power is the pretest-posttest design with a historical control or comparison group. However, this design may have questionable validity because of the historical control group. Can the researcher

demonstrate that the historical control group is comparable to the experimental group? The repeated measures design increases power if the trait being assessed is relatively stable over time. Designs that use blocking or stratification usually require an increase in the total sample size. The sample size increases in proportion to the number of cells included in the data analysis. Designs that use matched pairs of subjects have greater power and require a smaller sample (see Chapter 11 for a discussion of these designs). The higher the degree of correlation between subjects on the variable on which the subjects are matched, the greater the power (Kraemer & Thiemann, 1987).

Kraemer and Thiemann (1987) classified studies as *exploratory* or *confirmatory*. According to their approach, confirmatory studies should be conducted only after a large body of knowledge has been gathered through exploratory studies. **Confirmatory studies** are expected to have large samples and to use random sampling techniques. These expectations are lessened for exploratory studies. **Exploratory studies** are not intended for generalization to large populations. They are designed to increase the knowledge of the field of study. For example, pilot or preliminary studies to test a methodology or provide estimates of an *ES* are often conducted before a larger study. In other studies, the variables, not the subjects, are the primary area of concern. Several studies may examine the same variables using different populations. In these types of studies, the specific population used may be incidental. Data from these studies may be used to define population parameters. This information can be used to conduct confirmatory studies using large, randomly selected samples.

Confirmatory studies, such as studies testing the effects of nursing interventions on patient outcomes or studies testing the fit of a theoretical model, require large sample sizes. Clinical trials are being conducted in nursing for these purposes. The power of these large, complex studies must be carefully analyzed (Leidy & Weissfeld, 1991). For the large sample sizes to be obtained, subjects are acquired in numerous clinical settings, sometimes in different parts of the United States. Kraemer and Thiemann (1987) believed that these studies should not be performed until extensive information is available from exploratory studies. This information should include meta-analysis and the definition of a population *ES*.

Number of Variables

As the number of variables under study grows, the needed sample size may also increase. Adding variables such as age, gender, ethnicity, and education to the analysis plan (just to be on the safe side) can increase the sample size by a factor of 5 to 10 if the selected variables are uncorrelated with the dependent variable. In this case, instead of a sample of 50, you may need a sample of 250 to 500 if you plan to use the variables in the statistical analyses. (Using them only to describe the sample does not cause a problem in terms of power.) If the variables are highly correlated with the dependent variable, however, the effect size will increase, and the sample size can be reduced.

Variables included in the data analysis must be carefully selected. They should be essential to the research purpose or should have a documented strong relationship with the dependent variable (Kraemer & Thiemann, 1987). Sometimes researchers have obtained sufficient sample size for the primary analyses but failed to plan for analyses involving subgroups, such as analyzing the data by age categories or by ethnic groups, which require a larger sample size. A larger sample size is also needed if multiple dependent variables have been included.

Measurement Sensitivity

Well-developed instruments measure phenomena with precision. For example, a thermometer measures body temperature precisely. Instruments measuring psychosocial variables tend to be less precise. However, a scale with strong reliability and validity tends to measure more precisely than an instrument that is less well developed. Variance tends to be higher in a less well-developed tool than in one that is well developed. An instrument with a smaller variance is preferred because the power of a test always decreases when within-group variance increases (Kraemer & Thiemann, 1987). If you were measuring anxiety and the actual anxiety score for several subjects was 80, the subjects' scores on a less well-developed scale might range from 70 to 90, whereas a well-developed scale would tend to show a score closer to the actual score of 80 for each subject. As variance in instrument scores increases, the sample size needed to gain an accurate understanding of the phenomenon under study increases.

The range of measured values influences power. For example, a variable might be measured in 10 equally spaced values, ranging from 0 to 9. *ESs* vary according to how near the value is to the population mean. If the mean value is 5, *ESs* are much larger in the extreme values and lower for values near the mean. If you decided to use only subjects with values of 0 and 9, the *ES* would be large, and the sample could be small. The credibility of the study might be questionable, however, because the values of most individuals

would not be 0 or 9 but rather would tend to be in the middle range of values. If you decided to include subjects who have values in the range of 3 to 6, excluding the extreme scores, the *ES* would be small, and you would require a much larger sample. The wider the range of values sampled, the larger the *ES* (Kraemer & Thiemann, 1987). If you had a heterogeneous group of study participants, you would expect them to have a wide range of scores on a depression scale, which would increase the *ES*. A strong measurement method has validity and reliability and measures variables at the interval or ratio level. The stronger the measurement methods used in a study, the smaller the sample that is needed to identify significant relationships among variables and differences between groups.

Data Analysis Techniques

Data analysis techniques vary in their ability to detect differences in the data. Statisticians refer to this as the power of the statistical analysis. For your data analysis, choose the most powerful statistical test appropriate to the data. Overall, parametric statistical analyses are more powerful than nonparametric techniques in detecting differences and should be used if the data meet criteria for parametric analysis. However, in many cases, nonparametric techniques are more powerful if your data do not meet the assumptions of parametric techniques. Parametric techniques vary widely in their capacity to distinguish fine differences and relationships in the data. Parametric and nonparametric analyses are discussed in Chapter 21.

There is also an interaction between the measurement sensitivity and the power of the data analysis technique. The power of the analysis technique increases as precision in measurement increases. Larger samples must be used when the power of the planned statistical analysis is low.

For some statistical procedures, such as the *t*-test and analysis of variance (ANOVA), having equal group sizes increases power because the effect size is maximized. The more unequal the group sizes are, the smaller the effect size. In unequal groups, the total sample size must be larger (Kraemer & Thiemann, 1987).

The chi-square (χ^2) test is the weakest of the statistical tests and requires very large sample sizes to achieve acceptable levels of power. As the number of categories (cells in the chi-square analysis) in a study grows, the sample size needed increases. Also, if there are small numbers in some of the categories, you must increase the sample size. Kraemer and Thiemann (1987) recommended that the chi-square test be used only when no other options are available. In addition, the categories should be limited to those essential to the study.

Sample Size in Qualitative Research

In quantitative research, the sample size must be large enough to describe variables, identify relationships among variables, or determine differences between groups. However, in qualitative research, the focus is on the quality of information obtained from the person, situation, event, or documents sampled versus the size of the sample (Marshall & Rossman, 2011; Munhall, 2012; Patton, 2002; Sandelowski, 1995). The sample size and sampling plan are determined by the purpose and philosophical basis of the study. The sample size required is determined by the depth of information needed to gain insight into a phenomenon, explore and describe a concept, describe a cultural element, develop a theory, or describe a historical event. The sample size can be too small when the data collected lack adequate depth or richness. An inadequate sample size can reduce the quality and credibility of the research findings. Many qualitative researchers use purposive or purposeful sampling methods to select the specific participants, events, or situations that they believe would provide them the rich data needed to gain insights and discover new meaning in an area of study (Sandelowski, 2000).

The adequacy of the sample size in a study should be justified by the researchers. Often the number of participants in a qualitative study is adequate when saturation of information is achieved in the study area (Fawcett & Garity, 2009). **Saturation of data** occurs when additional sampling provides no new information, only redundancy of previously collected data. Important factors that must be considered in determining sample size to achieve saturation of data are (1) scope of the study, (2) nature of the topic, (3) quality of the data, and (4) study design (Marshall & Rossman, 2011; Morse, 2000; Munhall, 2012; Patton, 2002).

Scope of the Study

If the scope of a study is broad, researchers need extensive data to address the study purpose, and it takes longer to reach data saturation. A study with a purpose that has a broad scope requires more sampling of participants, events, or documents than a study with a narrow scope (Morse, 2000). A study that has a clear focus and provides focused data collection usually has richer, more credible findings. The depth of a study's scope and its clarity of focus influence the number

of participants needed for the study sample. For example, fewer participants would be needed to describe the phenomenon of chronic pain in adults with rheumatoid arthritis than would be needed to describe the phenomenon of chronic pain in elderly adults. A study of chronic pain experienced by elderly adults has a much broader focus, with less clarity, than a study of chronic pain experienced by adults with a specific medical diagnosis of rheumatoid arthritis.

Nature of the Topic

If the topic of your study is clear and the participants can easily discuss it, fewer individuals are needed to obtain the essential data. If the topic is difficult to define and awkward for people to discuss, you will probably need a larger number of participants or informants to saturate the data (Morse, 2000; Patton, 2002). For example, a phenomenological study of the experience of an adult living with a history of child sexual abuse is a sensitive, complex topic to investigate. This type of topic would probably require a greater number of participants and increased interview time to collect the essential data.

Quality of the Data

The quality of information obtained from an interview, observation, or document review influences the sample size. The higher the quality and richness of the data, the fewer the research participants needed to saturate data in the area of study. Quality data are best obtained from articulate, well-informed, and communicative participants. These participants are able to share more rich data in a clear and concise manner. In addition, participants who have more time to be interviewed usually provide data with greater depth and breadth. Qualitative studies require that you critically appraise the quality of the richness of communication elicited from the participants, the degree of access provided to events in a culture, or the number and quality of documents studied. These characteristics directly affect the richness of the data collected and influence the sample size needed to achieve quality study findings (Fawcett & Garity, 2009; Munhall, 2012).

Study Design

Some studies are designed to increase the number of interviews with participants. The more interviews conducted with a participant, the greater the quality of the data collected. For example, a study design that includes an interview both before and after an event would produce more data than a single interview. Designs that involve interviewing a family or a group of individuals produce more data than an interview with a single study participant. In critically appraising a qualitative study, determine if the sample size is adequate for the design of the study.

Beaulieu et al. (2011) provided a detailed discussion of how they determined their final sample size. This qualitative study conducted with grounded theory methodology was introduced earlier in the discussion of theoretical sampling. The study focused on developing a theory about decision making of young adult couples regarding their use of ECPs. The sample was obtained with convenience, snowball, and theoretical sampling and resulted in a sample size of 22 couples. The following study excerpt provides the researchers' rationale for the final sample size of their study.

> "A convenience sample was recruited via public notices and snowball sampling.… All interested young women initiated the first contact with the researcher by e-mail or telephone.… At the first meeting, which also included partners, study procedures were reviewed with participants, after which written consent and demographic information were obtained.…
>
> Analysis began simultaneously with data collection as dictated by the tenets of grounded theory.… As these processes progressed, axial coding was performed to identify core categories and their relationships. As new categories emerged, the original interview guide was revised and additional couples were recruited to allow for theoretical sampling, that is sampling specifically to fill in theoretical gaps, strengthen categories and their relationships, and verify or challenge emerging conceptualizations (Strauss & Corbin, 1998) and focused coding (Charmaz, 2006). Member checking occurred throughout the analysis by sharing the preliminary findings with subsequent couples to meet the requirements of confirmability of developing conceptualizations (Denzin & Lincoln, 2000). Saturation—when no new categories emerge (Strauss & Corbin, 1998)—was reached after interviewing 18 couples, but five more couples were included to ensure comprehensive analysis as well as theoretical verification. As the analysis continued through the processes of grounded theorizing, salient categories consistent with contemporary grounded theory principles (Clarke, 2005) were constructed to characterize the experience of young couples regarding ECPs. A theoretical model was developed to describe and explain the process of emergency contraceptive decision making in young couples." (Beaulieu et al., 2011, p. 43)

The study by Beaulieu et al. (2011) has many strengths in the area of sampling, including quality sampling methods (convenience, snowball, and theoretical), strong sample size (N = 22 couples), and conscientious participants. The investigators provide extensive details of the theoretical sampling conducted to ensure saturation was achieved with no new categories emerging. The saturation occurred after 18 couples, but the researchers interviewed 5 more couples to ensure depth and breadth in the data for theoretical verification. Beaulieu et al. (2011) described how they were able to develop successfully a theoretical model of experiences of young couples regarding use of ECPs. The study would have been strengthened by knowing how many study participants were obtained by each of the sampling methods (convenience, snowball, and theoretical). Also the researchers mentioned saturation was obtained with 18 couples but 5 more couples were included or N = 23 but the sample size identified was N = 22. A rationale is needed for the attrition of one of the couples from the study.

Research Settings

The **setting** is the location where a study is conducted. There are three common settings for conducting nursing research: natural, partially controlled, and highly controlled. A **natural setting**, or *field setting,* is an uncontrolled, real-life situation or environment (Kerlinger & Lee, 2000). Conducting a study in a natural setting means that the researcher does not manipulate or change the environment for the study. Descriptive and correlational quantitative studies and qualitative studies are often conducted in natural settings. For example, in the study by Beaulieu et al. (2011) discussed previously, the investigators made no attempt to manipulate the settings when they conducted the interviews for their qualitative study. All the data were conducted in natural settings as the interviews of the couples took place in various public settings or in their homes. The intent of the study was to develop a theory of the decision making of young adult couples regarding ECP use in a natural environment.

A **partially controlled setting** is an environment that the researcher manipulates or modifies in some way. An increasing number of nursing studies, usually correlational, quasi-experimental, and experimental studies, are being conducted in partially controlled settings. In a study that was introduced earlier in the discussion of sampling criteria, Twiss et al. (2009) conducted a quasi-experimental study to determine the effects of a strength training (ST) intervention on muscle strength, balance, and falls of breast cancer survivors (BCSs) with bone loss. These researchers used both partially controlled and natural settings in their studies, which are described in the following excerpt.

> "This was a multisite, randomized controlled trial of a 24-month multi-component intervention with follow-up data collection at 36 months.…
>
> **"Setting**
>
> "Exercise activities were performed in participants' homes [natural setting] or at investigator-approved fitness centers [partially controlled setting], and Biodex and balance testing were performed by physical therapists at hospitals or rehabilitation centers [partially controlled settings] at each of the four sites." (Twiss et al., 2009, p. 22)

Twiss et al. (2009) partially controlled the fitness centers by approving them but did not try to control the specific activities of each center. The researchers also ensured that the measurements were taken in a precise and accurate way by an expert (physical therapist) in partially controlled settings of hospitals and rehabilitation centers. The natural and partially controlled settings seemed appropriate in this study for the implementation of the ST intervention and for the precise and accurate measurement of the outcome variables.

A **highly controlled setting** is an artificially constructed environment developed for the sole purpose of conducting research. Laboratories, research or experimental centers, and test units in hospitals or other healthcare agencies are highly controlled settings where experimental studies are often conducted. This type of setting reduces the influence of extraneous variables, which enables the researcher to examine accurately the effect of one variable on another. Highly controlled settings are commonly used to conduct experimental research. Sharma, Ryals, Gajewski, and Wright (2010) conducted an experimental study to examine the effects of a moderate-intensity aerobic exercise program on painlike behavior and neurotrophin-3 (NT-3) in female mice. The rationale for conducting this study was that the literature and clinical practice supported the use of aerobic exercise in reducing pain and improving function in people with chronic pain, but the molecular basis for these positive actions was poorly understood. This study was conducted in a laboratory using a selected type of

mouse, and the setting is described in the following excerpt.

> "All experiments were approved by the Institutional Animal Care and Use Committee of the University of Kansas Medical Center and adhered to the university's animal care guidelines. Forty CF-1 female mice (weight = 25 g) were used to examine the effects of moderately intense exercise on primary (muscular) and secondary (cutaneous) hyperalgesia and NT-3 synthesis. Because women develop wide-spread pain syndromes at a greater rate than age-matched men, hyperalgesia was induced in female mice. The mice were exposed to 12-hour light/dark cycle and had access to food and water *ad libitum*.…
>
> Initially, the mice were randomly assigned to either the acidic saline injection (experimental) group or the normal saline injection (placebo) group. Five days after inducing hyperalgesia with acidic saline injection into the right limb, the animals were further assigned to either exercise or no-exercise group.… Two 6-lane, motorized treadmills were used for the exercise training." (Sharma et al., 2010, pp. 715-716)

Sharma et al. (2010) conducted their study in a highly controlled laboratory setting in terms of the housing and feeding of the mice, the light and temperature of the environment, implementation of the treatments, and the measurements of the dependent variables. Only with animals could this type of setting control be achieved in conducting this study. This type of highly controlled setting removes the effect of numerous extraneous variables, so the effects of the independent variables on the dependent variables can be clearly determined. However, because this research was conducted on animals, the findings cannot be generalized to humans, and additional research is needed to determine the molecular basis of the influence of aerobic exercise on pain and functioning in humans.

Recruiting and Retaining Research Participants

After a research team makes a decision about the size of the sample, the next step is to develop a plan for **recruiting research participants**, which involves identifying, accessing, and communicating with potential study participants who are representative of the target population. Recruitment strategies differ, depending on the type of study, population, and setting. Special attention must focus on recruiting subjects who tend to be underrepresented in studies, such as minorities, women, children, and elderly adults (Fulmer, 2001; Gul & Ali, 2010; Hendrickson, 2007; Hines-Martin, Speck, Stetson, & Looney, 2009). The sampling plan, initiated at the beginning of data collection, is almost always more difficult than expected. In addition to subject recruitment, retaining acquired subjects is critical to achieve an acceptable sample size and requires researchers to consider the effects of the data collection strategies on subject attrition. **Retaining research participants** involves the participants or subjects completing the required behaviors of a study to its conclusion. The problems with retaining participants increase as the data collection period lengthens. Some researchers never obtain their planned sample size because of the problems they encounter as they try to recruit and retain subjects. These researchers often are forced to complete their study with a smaller sample size, which could decrease the power of the study and potentially produce nonsignificant results (Aberson, 2010). With an increasing number of studies being conducted in health care, recruiting and retaining subjects have become more complex issues for researchers to manage (Gul & Ali, 2010; McGregor, Parker, LeBlanc, & King, 2010).

Recruiting Research Participants

The effective recruitment of subjects is crucial to the success of a study. A few studies examining the effectiveness of various strategies of participant recruitment and retention have appeared in the literature (Davidson, Cronk, Harrar, Catley, & Good, 2010; Hines-Martin et al., 2009; Whitebird, Bliss, Savik, Lowry, & Jung, 2010). However, most of the information available to guide researchers comes from the personal experiences of skilled researchers, some of whom have published their ideas (Gul & Ali, 2010; McGregor et al., 2010). Some of the positive and negative factors that influence a subject's decision to participate in a study are (1) the attitudes and ethics of the researchers, (2) the subject's need for a treatment, (3) the subject's interest in the study topic, (4) financial compensation, (5) fear of the unknown, (6) time and travel constraints, (7) language barriers, and (8) the nature of the informed consent (Gul & Ali, 2010; Hine-Martin et al., 2009; Madsen et al., 2002; Papadopoulos & Lees, 2002; Sullivan-Bolyai et al., 2007).

The researcher's initial communication with a potential subject usually strongly affects the subject's decision about participating in the study. Therefore, the approach must be pleasant, positive, informative, culturally sensitive, and nonaggressive. The researcher

needs to explain the importance of the study and clarify exactly what the subject will be asked to do, how much of the subject's time will be involved, and what the duration of the study will be. Research participants are valuable resources, and the researcher must communicate this value to the potential subject. High-pressure techniques, such as insisting that the subject make an instant decision to participate in a study, usually lead to resistance and a higher rate of refusals. If the study involves minorities, researchers must be culturally competent or knowledgeable and skilled in relating to the particular ethnic group being studied (Hines-Martin et al., 2009; Papadopoulos & Lees, 2002). If the researcher is not of the same culture as the potential subjects, he or she may employ a data collector who is of the same culture. Hendrickson (2007) used a video for recruiting Hispanic women for her study, and she provided all the details related to the study in the subjects' own language in the video. This approach greatly improved the subjects' understanding of the study and their desire to participate.

If a potential subject refuses to participate in a study, you must accept the refusal gracefully—in terms of body language as well as words. Your actions can influence the decision of other potential subjects who observe or hear about the encounter. Studies in which a high proportion of individuals refuse to participate have a serious validity problem. The sample is likely to be biased because often only a certain type of individual has agreed to participate. You should keep records of the numbers of persons who refuse and, if possible, their reasons for refusal. With this information, you can include the refusal rate in the published research report with the reasons for refusal. It would also be helpful if you could determine if the potential subjects who refused to participate differed from the individuals who agreed to participate in the study. This information will help you to determine the representativeness of your sample (Thompson, 2002).

Recruiting minority subjects for a study can be particularly problematic. Minority individuals may be difficult to locate and are often reluctant to participate in studies because of feelings of being "used" while receiving no personal benefit from their involvement or because of their distrust of the medical community. Effective strategies for recruiting minorities include developing partnerships with target groups, community leaders, and potential participants in the community; using active face-to-face recruitment in nonthreatening settings; and using appropriate language to communicate clearly the purpose, benefits, and risks of the study (Alvarez, Vasquez, Mayorga, Feaster, & Mitrani, 2006). Hines-Martin et al. (2009) studied the recruitment and retention process for intervention research conducted with a sample of primarily low-income African American women. Their complex, multistage recruitment strategies are introduced in the following excerpt.

> "Phase I involved the development of a recruitment team, composed of a co-investigator, in addition to an African American nurse familiar with the target population, and two women who were long-standing community members.
>
> Phase I activities began with periods of observation in the community setting and discussions with community center personnel to improve the investigators' understanding of who used the community center services and when. It became increasingly clear that only two of the three communities felt a connection with or used the community center routinely.… Therefore, the recruitment team, with the assistance from nursing graduate students, walked every block of the two relevant communities at different times of the day and different days of the week to better understand when and where community women could be found in their daily lives.… Community women were informed of new initiatives at the center and were provided with recruitment flyers including pictures of the research team. The recruitment team then undertook usual recruitment activities, such as meeting with women's groups in the communities and recruitment at community fairs." (Hines-Martin et al., 2009, pp. 665-666)

If researchers use data collectors in their studies, they need to verify the data collectors are following the sampling plan, especially in random samples. When the data collectors encounter difficult subjects or are unable to make contact easily, they may simply shift to the next person without informing the principal investigator. This behavior could violate the rules of random sampling and bias the sample. If data collectors do not understand, or do not believe in, the importance of randomization, their decisions and actions can undermine the intent of the sampling plan. Thus, data collectors must be carefully selected and thoroughly trained. A plan for the supervision and follow-up of data collectors to increase their accountability should be developed (Thompson, 2002).

If you conduct a survey study, you may never have personal contact with the subjects. To recruit such subjects, you must rely on the use of attention-getting techniques, persuasively written material, and

strategies for following up on individuals who do not respond to the initial written or email communication. The strategies need to be appropriate to the potential subjects; mailed surveys are probably still the best way to obtain information from elderly adults. Because of the serious problems of low response rates in survey studies, using strategies to increase the response rate is critical. For instance, we have received a teabag or packet of instant coffee with a questionnaire, accompanied by a recommendation in the letter to have a cup of tea or coffee "on" the researcher while we complete the questionnaire. Creativity is required in the use of such strategies because they tend to lose their effect on groups who receive questionnaires frequently. In some cases, small amounts of money ($1.00 to $5.00) are enclosed with the letter, which may suggest that the recipient buy a soft drink or that the money is a small offering for completing the questionnaire. This strategy imposes some sense of obligation on the recipient to complete the questionnaire, but it is not thought to be coercive. Also, you should plan emailing or mailings to avoid holidays or times of the year when activities are high for potential subjects, possibly reducing the return rate.

Researchers frequently use the Internet to recruit subjects and to collect survey data. This method makes it easier for you to contact potential subjects and for the subjects to provide the requested data. However, an increased number of surveys are being sent by the Internet, which can decrease the response rate of potential subjects who are frequently surveyed. In studies with surveys, the letter emailed to potential subjects must be carefully composed. It may be your only chance to persuade the subject to invest the time needed to complete the questionnaire. You must sell the reader on the importance of both your study and his or her response. The tone of your letter will be the potential subject's only image of you as a person; yet, for many subjects, their response to the perception of you as a person most influences their decision about completing the questionnaire. Seek examples of letters sent by researchers who have had high response rates, and save letters you received to which you responded positively. You also might pilot-test your letter on potential research participants who can give you feedback about their reactions to the letter's tone.

The use of follow-up emails, letters, or cards has been repeatedly shown to raise response rates to surveys. The timing is important. If too long a period has lapsed, the potential subject may have deleted the questionnaire from his or her email box or discarded the mailed copy. However, sending the follow-up too soon could be offensive. Before the questionnaires are emailed or mailed, precise plans need to be made for monitoring the return of each questionnaire. A bar graph could be developed to record the return of each questionnaire as a means of suggesting when the follow-up mailing or emailing should occur. The cumulative number and percentage of responses would be logged on the graph to reflect the overall data collection process. The data from emailed questionnaires can be immediately analyzed so that researchers can easily keep track of the numbers of participants responding. When the daily or weekly responses decline, a follow-up email or first-class letter could be sent encouraging individuals to complete the questionnaire. Study participants and questionnaires are assigned the same code numbers, and nonrespondents are identified by checking the list of code numbers of unreturned questionnaires. A third follow-up questionnaire with a further modified cover letter could be emailed or mailed to increase the return rate for the questionnaires.

The factors involved in the decision of whether to respond to a questionnaire are not well understood. One factor is the time required to respond; this includes the time needed to orient to the directions and the emotional energy necessary to deal with the threats and anxieties generated by the questions. There is also a cognitive demand for thinking. Subjects seem to make a judgment about the relevance of the research topic and the potential for personal application of findings. Previous experience with questionnaires is also a deciding factor.

Traditionally, subjects for physiological nursing studies have been sought in the hospital setting. However, access to these subjects is becoming more difficult—in part because of the larger numbers of nurses and other healthcare professionals now conducting research. The largest involvement of research subjects within a healthcare agency usually occurs with medical research and mainly with clinical trials that include large samples (Gul & Ali, 2010). Nurse researchers are recruiting subjects from a variety of clinical settings. Whitebird et al. (2010) identified three successful recruitment methods to use in healthcare agencies: (1) identifying potential participants using administrative databases, (2) obtaining referrals of potential participants through healthcare providers and other sources, and (3) approaching directly a known potential subject. An initial phase of recruitment may involve obtaining community and institutional support for the study. Support from other healthcare professionals, such as nurses and physicians, and clinical agency staff is usually crucial to the successful recruitment of research participants.

Recruitment of subjects for clinical trials requires a different set of strategies because the recruitment may be occurring simultaneously in several sites (perhaps in different cities). Many of these studies never achieve their planned sample size. The number of subjects meeting the sampling criteria who are available in the selected clinical sites may not be as large as anticipated. Researchers must often screen twice as many patients as are required for the study to obtain a sufficient sample size. Screening logs must be kept during the recruiting period to record data on patients who met the criteria but were not entered into the study. Researchers commonly underestimate the amount of time required to recruit subjects for a clinical trial. In addition to defining the number of subjects and the time set aside for recruitment, it may be helpful to develop short-term or interim recruitment goals designed to maintain a constant rate of patient entry (Gul & Ali, 2010). Hellard, Sinclair, Forbes, and Fairley (2001) studied methods to improve the recruitment and retention of subjects in clinical trials and found that the four most important strategies were to (1) use nonaggressive recruitment methods, (2) maintain regular contact with the participants, (3) ensure that the participants are kept well informed of the progress of the study, and (4) provide constant encouragement to subjects to continue participation. Sullivan-Bolyai et al. (2007) identified the barriers and strategies to improve the recruitment of study participants from clinical settings. Table 15-3 identifies these common barriers to research participant recruitment and provides possible strategies to manage them.

Studies may also benefit from the endorsement of community leaders, such as city officials; key civic leaders; and leaders of social, educational, religious, or labor groups. In some cases, these groups may be involved in planning the study, leading to a sense of community ownership of the project. Community groups may also help researchers to recruit subjects for the study. Subjects who meet the sampling criteria sometimes are found in the groups assisting with the study. Endorsement may involve letters of support and, in some cases, funding. These activities can add legitimacy to the study and make involvement in the study more attractive to potential subjects (Alvarez et al., 2006; Davidson et al., 2010; Hines-Martin et al., 2009).

Media support can be helpful in recruiting subjects. Researchers can place advertisements in local newspapers and church and neighborhood bulletins. Radio stations can make public service announcements. Members of the research team can speak to groups relevant to the study population. Your team can place posters in public places, such as supermarkets, drugstores, and public laundries. With permission, you can set up tables in shopping malls with a member of the research team present to recruit subjects. Plan for possible challenges in recruitment and include multiple methods and locations in your application for human subject approval for your study. Otherwise, you would need to submit a modified protocol to the institutional review board when you add a method or site for recruitment.

Davidson et al. (2010) used multiple strategies to recruit and retain college smokers in a cessation clinical trial. Their four-phase recruitment process is presented in the following study excerpt.

> "Participants in this study were members of Greek fraternities and sororities enrolled at a large Midwestern university, and data were collected from 2006 to 2009.… The clinical trial involved testing a four-session, MI [motivational interviewing] counseling intervention on smoking cessation. Participants were recruited from college fraternity and sorority chapters regardless of their interest in quitting smoking. Recruitment involved four phases. First, out of 41 fraternity and sorority chapters from a large Midwestern university, the 30 chapters with the larger memberships were invited to participate. Second, within these invited chapters, individuals were recruited to participate in an initial, 5-minute, 8-item screening survey (i.e., Screener).
>
> Third, individual members of these 30 chapters who met the inclusion criteria based on the Screener and who were interested in participating in the study were recruited to participate in a more extensive (30-45 minute) computerized baseline assessment approximately 1-4 days following the Screener.… Fourthly, eligible individuals who completed the baseline assessment were recruited for enrollment in the clinical trial." (Davidson et al., 2010, pp. 146-147)

The recruitment for this smoking cessation clinical trial was accomplished by using the Greek chapters. Davidson et al. (2010) developed relationships with these Greek organizations by meeting with leaders and members and attending special events. To accomplish phases two and three, the researchers met with the participants at convenient times and in accessible locations. The participants were also provided incentives of food (cookies and pizza), small cash gifts, and raffles for iPods. These creative strategies increased the recruitment and retention of the study participants.

TABLE 15-3 Barriers to Recruitment with Actions and Strategies for Engaging Health Care Providers in the Referral Process

Barriers and Actions	Strategies
HIPAA*	Ask Clinicians to distribute letters to potential study participants
Create alternative recruitment methods	Obtain institutional review board waiver of authorization requirement for the use or disclosure of personal health information
	Work with clinics to secure a consent that meets HIPAA* regulations and allows the staff to provide names and contact information of patients with specific conditions that may be of interest to researchers
	Recognize and acknowledge the burden that recruitment places on healthcare providers
Work burden	Provide Salary support
Create compensations	Provide educational incentives (e.g., purchase laptop, journals, books, pay for conference attendance in the field under study) for healthcare providers who do not normally have access to such opportunities as part of their job
	Assess administrative or managerial perceptions of healthcare providers' recruitment-related responsibilities, and if salary support is given, how that money will be used
	Discuss the designated recruitment tasks and responsibilities with the assigned staff to determine their perceptions and expectations
Financial disincentives	Assess the clinic's financial situation and determine if it is realistic, pragmatic, or feasible to use that site, especially if its funding depends on patient numbers
Recognize that patient numbers or productivity may be linked to the clinic's livelihood	Help keep participants linked to the clinical site while they are participating in the study
Provider competition	Develop a research proposal that reflects the clinical site's philosophical and policy perspectives and priorities
Create a partnership with healthcare providers involved in recruitment so that they are rewarded and acknowledged for their participation in the research process	Include healthcare providers in the development of a study
	Hire and pay a clinical staff member to be responsible for introducing the study to potential participants
	Link recruitment activities to nursing clinical ladder or organization values
	Maintain open communication between the clinical and research teams regarding the workings of the study
Provider concerns	Assess healthcare providers' perceptions of research
Demystify research process	Encourage healthcare providers to participate in developing the research proposal
Develop a team atmosphere and a spirit of "we're all in this together"	Include healthcare providers in developing study-related manuscripts
	Include healthcare providers in research team meetings at a mutually convenient time
	Express appreciation in an ongoing basis for healthcare providers' involvement in recruitment process
	Share recruitment status information on a monthly basis with healthcare providers
	Share pilot or feasibility data with healthcare providers to support the study rationale and choice of specific methods
Desire to protect patients	Acknowledge responsibility of healthcare providers to protect patients from harm
Work with healthcare providers to acknowledge and respect patient decision-making abilities	Address concerns of healthcare providers by emphasizing the pilot data that supports the protocol
Encourage healthy partnerships between patients and healthcare providers	Model respectful partnerships with study participants

From Sullivan-Bolyai, S., Bova, C., Deatrick, J. A., Knafl, K., Grey, M., Leung, K., et al. (2007). Barriers and strategies for recruiting study participants in clinical settings. *Western Journal of Nursing Research, 29*(4), 498–499.

**HIPAA,* Health Insurance Portability and Accountability Act.

Retaining Subjects

A serious problem in many studies is subject retention, and sometimes participant attrition cannot be avoided. Subjects move, die, or withdraw from a treatment. If you must collect data at several points over time, subject attrition can become a problem. Subjects who move frequently and subjects without phones pose a particular problem. Numerous strategies have been found to be effective in maintaining the sample. It is a good idea to obtain the names, email addresses, and phone numbers (cell and home numbers if possible) of at least two family members or friends when you enroll the participant in the study. Ask if the participant would agree to give you access to unlisted phone numbers in the event of changes in his or her number.

In some studies, subjects are reimbursed for time and expenses related to participation. A bonus payment may be included for completing a certain phase of the study. Gifts can be used in place of money. Sending greeting cards for birthdays and holidays helps maintain contact. Researchers found that money was more effective than gifts in retaining subjects in longitudinal studies. However, some people pointed out the moral issues related to providing monetary payment to subjects. This strategy can compromise the voluntariness of participation in a study and particularly has the potential of exploiting low-income persons.

Collecting data takes time. The participant's time is valuable and should be used frugally. During data collection, it is easy to begin taking the participant for granted. Taking time for social amenities with participants may also pay off. However, take care that these interactions do not influence the data being collected. Beyond that, nurturing subjects participating in the study is critical. In some situations, providing refreshments and pleasant surroundings is helpful. During the data collection phase, you also may need to nurture others who interact with the participants; these may be volunteers, family, staff, students, or other professionals. It is important to maintain a pleasant climate for the data collection process, which pays off in the quality of data collected and the retention of subjects (Davidson et al., 2010; Gul & Ali, 2010; Hines-Martin et al., 2009; McGregor et al., 2010).

Qualitative studies and longitudinal studies require extensive time commitment from subjects. They are asked to participate in detailed interviews or to complete numerous forms at various intervals during a study (Marshall & Rossman, 2011; Munhall, 2012; Patton, 2002). Sometimes data are collected with diaries that require daily entries over a set period of time. These studies face the greatest risk of participant mortality. Chapters 4 and 12 provide more details on the recruitment and retention of research participants for qualitative studies.

Clinical trials can also require extensive time commitments from subjects. Gul and Ali (2010) mentioned the importance of overcoming participant barriers to continuing in a study, such as time to complete data collection forms, transportation problems, and conflicts with work and family commitments. There is no formula for compensating study participants, but many studies mention small monetary payments, gifts, or free health or child care. It is important that the incentives used to recruit and retain research participants be documented in the published study. Communication is one of the most important facets to retaining study participants. Davidson et al. (2010), whose recruitment strategies were introduced earlier, describe their success with retention in their smoking cessation clinical trial in the following excerpt.

> "A very high proportion of participants (89%) completed at least one session (90% treatment; 87% comparison). The majority (73%) were retained, completing three or more sessions (75% treatment; 70% comparison), and over half completed the maximum of four sessions (63% treatment; 61% comparison). At the follow-up assessment occurring 6 months after the baseline assessment, 79% of the participants ($n = 357$) were retained (80% treatment; 78% comparison)." (Davidson et al., 2010, p. 150)

Research participants who have a personal investment in the study are more likely to complete the study. This investment occurs through interactions with and nurturing by the researcher. A combination of the participant's personal belief in the significance of the study, the perceived altruistic motives of the researcher in conducting the study, the ethical actions of the researcher, and the nurturing support provided by the researcher during data collection can greatly diminish subject attrition (Hines-Martin et al., 2009; Madsen et al., 2002; McGregor et al., 2010). The recruitment and retention of subjects will continue to be significant challenges for researchers, and creative strategies are needed to manage these challenges.

KEY POINTS

- Sampling involves selecting a group of people, events, behaviors, or other elements with which to conduct a study. Sampling denotes the process of

making the selections; sample denotes the selected group of elements.

- A sampling plan is developed to increase representativeness, decrease systematic bias, and decrease the sampling error; there are two main types of sampling plans—probability and nonprobability.
- Sampling error includes random variation and systematic variation. Refusal and attrition rates are important to calculate in a study to determine potential systematic variation or bias.
- Four sampling designs have been developed to achieve probability or random sampling: simple random sampling, stratified random sampling, cluster sampling, and systematic sampling.
- In nonprobability (nonrandom) sampling, not every element of the population has an opportunity for selection in the sample. The five nonprobability sampling designs described in this textbook are (1) convenience sampling, (2) quota sampling, (3) purposive or purposeful sampling, (4) network or snowball sampling, and (5) theoretical sampling.
- In quantitative studies, sample size is best determined by a power analysis, which is calculated using the level of significance (usually $\alpha = 0.05$), standard power of 0.80 (80%), and effect size. Factors important to sample size in quantitative research include (1) type of study, (2) number of variables studied, (3) measurement sensitivity, and (4) data analysis techniques.
- The number of participants in a qualitative study is adequate when saturation of information is achieved in the study area, which occurs when additional sampling provides no new information, only redundancy of previously collected data. Important factors that must be considered in determining sample size to achieve saturation of data are (1) scope of the study, (2) nature of the topic, (3) quality of the data, and (4) study design.
- The three common settings for conducting nursing research are natural, partially controlled, and highly controlled. A natural setting, or field setting, is an uncontrolled, real-life situation or environment. A partially controlled setting is an environment that the researcher has manipulated or modified in some way. A highly controlled setting is an artificially constructed environment, such as a laboratory or research unit in a hospital, developed for the sole purpose of conducting research.
- Recruiting and retaining research participants have become significant challenges in research; some strategies to assist researchers with these challenges are provided so that their samples might be more representative of their target population.

REFERENCES

Aberson, C. L. (2010). *Applied power analysis for the behavioral sciences*. New York, NY: Routledge Taylor & Francis Group.

Alvarez, R. A., Vasquez, E., Mayorga, C. C., Feaster, D. J., & Mitrani, V. B. (2006). Increasing minority research participation through community organization outreach. *Western Journal of Nursing Research*, *28*(5), 541–560.

Barhyte, D. Y., Redman, B. K., & Neill, K. M. (1990). Population or sample: Design decision. *Nursing Research*, *39*(5), 309–310.

Beaulieu, R., Kools, S. M., Kennedy, H. P., & Humphreys, J. (2011). Young adult couples' decision making regarding emergency contraceptive pills. *Journal of Nursing Scholarship*, *43*(1), 41–48.

Botman, S. L., Moore, T. F., Moriarty, C. L. & Parsons, V. L. (2000). Design and estimation for the National Health Interview Survey, 1995-2004. *National Center for Health Statistics, Vital Health Statistics, Series 2, No. 130*, 1–32.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Clarke, A. (2005). *Situational analysis grounded theory after the postmodern turn*. Thousand Oaks, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Academic Press.

Coyne, I. T. (1997). Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries. *Journal of Advanced Nursing*, *26*(3), 623–630.

Davidson, M. M., Cronk, N. J., Harrar, S., Catley, D., & Good, G. E. (2010). Strategies to recruit and retain college smokers in cessation trials. *Research in Nursing & Health*, *33*(2), 144–155.

Degirmen, N., Ozerdogan, N., Sayiner, D., Kosgeroglu, N., & Ayranci, U. (2010). Effectiveness of foot and hand massage in post-cesarean pain control in a group of Turkish pregnant women. *Applied Nursing Research*, *23*(3), 153–158.

Denzin, N., & Lincoln, Y. (2000). *Handbook of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Djukic, M., Kovner, C., Budin, W. C., & Norman, R. (2010). Physical work environment: Testing an expanded model of job satisfaction in a sample of registered nurses. *Nursing Research*, *59*(6), 441–451.

Fain, J. (2004). *Reading, understanding, and applying nursing research* (2nd ed.). Philadelphia, PA: F. A. Davis Company.

Fawcett, J., & Garity, J. (2009). *Evaluating research for evidence-based nursing practice*. Philadelphia, PA: F. A. Davis.

Floyd, J.A. (1993). Systematic sampling: Theory and clinical methods. *Nursing Research*, *42*(5), 290–293.

Foulabakhsh, J. M., & Stommel, M. (2010). Gender, symptom experience, and use of complementary and alternative medicine practices among cancer survivors in the U.S. cancer population. *Oncology Nursing Forum*, *37*(1), E7–E15. Retrieved from *Oncology Nursing Forum* online.

Fulmer, T. (2001). Editorial: Recruiting older adults in our studies. *Applied Nursing Research*, *14*(2), 63.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: IL. Aldine.

Green, J., & Thorogood, N. (2004). *Qualitative methods for health research*. Thousand Oaks, CA: Sage Publications.

Grove, S. K. (2007). *Statistics for health care research: A practical workbook*. St. Louis, MO: Saunders.

Gul, R. B., & Ali, P. A. (2010). Clinical trials: The challenge of recruitment and retention of participants. *Journal of Clinical Nursing*, *19*(1-2), 227–233.

Hellard, M.E., Sinclair, M.I., Forbes, A.B., & Fairley, C.K. (2001). Methods used to maintain a high level of participant involvement in a clinical trial. *Journal of Epidemiology & Community Health*, *55*(5), 348–351.

Hendrickson, S. G. (2007). Video recruitment of non-English-speaking participants. *Western Journal of Nursing Research*, *29*(2), 232–242.

Hines-Martin, V., Speck, B. J., Stetson, B., & Looney, S. W. (2009). Understanding systems and rhythms for minority recruitment in intervention research. *Research in Nursing & Health*, *32*(6), 657–670.

Jones, D., Duffy, M. E., & Flanagan, J. (2011). Randomized clinical trial testing efficacy of a nurse-coached intervention in arthroscopy patients. *Nursing Research*, *60*(2), 92–99.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth, TX: Harcourt College Publishers.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Newbury Park, CA: Sage.

Larson, E. (1994). Exclusion of certain groups from clinical research. *Image: Journal of Nursing Scholarship*, *26*(3), 185–190.

Leidy, N. K., & Weissfeld, L. A. (1991). Sample sizes and power computation for clinical intervention trials. *Western Journal of Nursing Research*, *13*(1), 138–144.

Levy, P. S., & Lemsbow, S. (1980). *Sampling for health professionals*. Belmont, CA: Lifetime Learning.

Li, Y., & Mukamel, D. B. (2010). Racial disparities in receipt of influenza and pneumococcus vaccinations among US nursing-home residents. *American Journal of Public Health*, *100*(S1), S256–S262.

Madsen, S. M., Mirza, M. R., Holm, S., Hilsted, K. L., Kampmann, K., & Riis, P. (2002). Attitudes towards clinical research amongst participants and nonparticipants. *Journal of Internal Medicine*, *251*(2), 156–168.

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage Publications.

McGregor, L., Parker, K., LeBlanc, P., & King, K. M. (2010). Using social exchange theory to guide successful study recruitment and retention. *Nurse Researcher*, *17*(2), 74–82.

Melnyk, B. M., & Fineout-Overholt, E. (2011). *Evidence-based practice in nursing & healthcare: A guide to best practice* (2nd ed.). Philadelphia: PA: Lippincott Williams & Wilkins.

Morse, J. M. (2000). Determining sample size. *Qualitative Health Research*, *10*(1), 3–5.

Munhall, P. L. (2007). *Nursing research: A qualitative perspective* (4th ed.). Sudbury, MA: Jones & Bartlett.

Munhall, P. L. (2012). *Nursing research: A qualitative perspective* (5th ed.). Sudbury, MA: Jones & Bartlett.

Neumark, D. E., Stommel, M., Given, C. W., & Given, B. A. (2001). Brief report: Research design and subject characteristics predicting nonparticipation in panel survey of older families with cancer. *Nursing Research*, *50*(6), 363–368.

O'Shea, E. R., Wallace, M., Griffin, M. Q., & Fitzpatrick, J. J. (2011). The effect of an educational session on pediatric nurses' perspectives toward providing spiritual care. *Journal of Pediatric Nursing*, *26*(1), 34–43.

Papadopoulos, I., & Lees, S. (2002). Developing culturally competent researchers. *Journal of Advanced Nursing*, *37*(3), 258–264.

Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pieper, B., Templin, T. N., Kirsner, R. S., & Birk, T. J. (2009). Impact of injection drug use on distribution and severity of chronic venous disorders. *Wound Repair & Regeneration*, *17*(4), 485–491.

Pieper, B., Templin, T. N., Kirsner, R. S., & Birk, T. J. (2010). The impact of vascular leg disorders on physical activity in methadone-maintained adults. *Research in Nursing & Health*, *33*(5), 426–440.

Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, *18*(2), 179–183.

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, *23*(3), 246–255.

Sharma, N. K., Ryals, J. M., Gajewski, B. J., & Wrights, D. E. (2010). Aerobic exercise alters analgesia and neurotrophin-3 synthesis in an animal model of chronic widespread pain. *Physical Therapy*, *90*(5), 714–725.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research, techniques, and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Sternberg, R. M., & Barry, C. (2011). Transnational mothers crossing the border and bringing their health care needs. *Journal of Nursing Scholarship*, *43*(1), 64–71.

Sullivan-Bolyai, S., Bova, C., Deatrick, J. A., Knafl, K., Grey, M., Leung, K., et al. (2007). Barriers and strategies for recruiting study participants in clinical settings. *Western Journal of Nursing Research*, *29*(4), 486–500.

Thompson, S. K. (2002). *Sampling* (2nd ed.). New York, NY: John Wiley & Sons.

Twiss, J. J., Waltman, N. L., Berg, K., Ott, C. D., Gross, G. J., & Lindsey, A. M. (2009). An exercise intervention for breast cancer survivors with bone loss. *Journal of Nursing Scholarship*, *41*(1), 20–27.

Ulrich, C. M., Danis, M., Ratcliffe, S. J., Garrett-Mayer, E., Koziol, D., Soeken, K. L., et al. (2006). Ethical conflict in nurse practitioners and physician assistants in managed care. *Nursing Research*, *55*(6), 391–401.

Whitebird, R. R., Bliss, D. Z., Savik, K., Lowry, A., & Jung, H. G. (2010). Comparing community and specialty provider-based recruitment in a randomized clinical trial: Clinical trial in fecal incontinence. *Research in Nursing & Health*, *33*(6), 500–511.

Yarandi, H. N. (1991). Planning sample sizes: Comparison of factor level means. *Nursing Research*, *40*(1), 57–58.

Yates, F. (1981). *Sampling methods for censuses and surveys*. New York, NY: Macmillan.