Epidemiological Field Work in Population-Based Studies

I.10

Arlène Fink

Introduction 10.1

Field work in epidemiological studies consists of collecting data in natural and experimental settings to answer research questions or test hypotheses about the origins, distribution, and control of disease in populations. Field data can be collected directly and indirectly. Although direct data collection traditionally includes collecting biological samples such as blood and saliva, epidemiologists also collect data about the health of populations by contacting respondents through the telephone, mail, or online. To study a community's use of preventative health services (such as influenza vaccinations), for example, a team of epidemiologists can conduct in-person or telephone interviews or administer written, computer-assisted or online surveys. Indirect data collection includes reviewing written, oral, and visual records of respondents' thoughts and actions and observing them in their natural or experimental environment. To study the extent to which a health care system's medical providers adhere to recommended guidelines for preventative health care, for instance, a team of epidemiologists might review a sample of medical records to identify which preventative services were used and by whom. If the team were interested in understanding why preventative services were (or were not) used, it might review transcripts of audio or videotapes of selected physician and patient encounters.

This chapter focuses on providing practical tips on the spectrum of techniques epidemiologists can use in designing and administering reliable and non-biological field measures. Although the chapter focuses on direct data collection, many of the principles apply also to indirect data collection.

Asking for Information: What Are the Characteristics of Straightforward Questions and Responses? 10.2

Learning how to ask questions in written and spoken form is essential when collecting field data. A straightforward question asks for information in unambiguous way and extracts accurate and consistent data. Straightforward questions are purposeful, use correct grammar and syntax, and call for one thought at a time with mutually exclusive questions (Sudman and Bradburn 1982; Fink 2002).

Types of Questions 10.2.1

Purposeful Questions. Questions are purposeful when the respondent can readily identify the relationship between the intention of the question and the objectives of the study. If the objectives are to find out about the uses of health services, for instance, and some of the study's questions ask about education

or place of birth, an explanation is needed of the relationship between questions and objectives. For instance, the introduction, can say something like: "We plan to compare people with differing backgrounds in their use of health services."

Concrete Questions. A concrete question is precise and unambiguous. Adding a dimension of time can help make the question more concrete. For instance, rather than ask: "Has a physician *ever* told you that you have hypertension?" ask, "In the *past* 12 *months* has a physician told you that you have hypertension?"

Complete Sentences. Questions should always be stated as complete sentences. Complete sentences express one entire thought, as in Example 1.

Example 1. *Complete Sentences*
Poor: Place of birth?
Comment: Place of birth means different things to different people. I might give the city in which I was born, but you might tell the name of the country or hospital.
Better: Name the country in which you were born. ♦

Make sure that experts and a sample of potential respondents review all questions even if you are using an already existing and validated instruments. Respondents' reading levels and attention spans may vary across studies.

Open and Closed Questions. Questions can take two primary forms. When they require respondents to use their own words, they are called open or open-ended. When the answers or responses are preselected for the respondent, the question is termed closed. Both types of questions have advantages and limitations.

An open question is useful when the intricacies of an issue are still unknown, in getting unanticipated answers, and for describing the world as the respondent sees it – rather than as the questioner does. Also, some respondents prefer to state their views in their own words and may resent the questioner's preselected choices. Sometimes, when left to their own devices, respondents provide quotable material. The disadvantage is that unless the team includes a trained anthropologist or qualitative researcher, responses to open questions are often difficult to interpret and compare.

Some respondents prefer closed questions because they are either unwilling or unable to express themselves. Closed questions are more difficult to write than open ones, however, because the answers or response choices must be known in advance. But the results lend themselves more readily to statistical analysis and interpretation. This feature is particularly important in most epidemiological studies which often rely on relatively large numbers of responses and respondents. Also, because the respondent's expectations are more clearly spelled out in closed questions, the answers have a better chance of being more reliable or consistent over time. Example 2 shows a closed question.

Example 2. *A Closed Question*
How often during the past week were you irritable? Circle one.

	Please Circle One
Always or nearly always	1
Sometimes	2
Rarely or never	3

Types of Responses

10.2.2

The choices given to respondents for their answers may take three forms (Fink and Kosecoff 1998; McDowell and Newell 1996; Stewart and Ware 1992). The first is called nominal or categorical. (The two terms are sometimes used interchangeably.) Categorical or nominal choices have no numerical or preferential values. For example, asking respondents if they are male or female is the same as asking them to "name" themselves as belonging to one of two categories: male or female.

The second form of response choice is called ordinal. When respondents are asked to rate or order choices, say, from very positive to very negative, they are given ordinal choices. The third form of response choice results in numerical data such as when a respondent is asked to give his or her height or age at the last birthday.

A hypothetical study finding that uses nominal data results in numbers or percentages, as follows:

> Five hundred respondents were interviewed in a study of preventative health. All were asked to indicate whether or not they perform four health-related activities: (1) exercise at least 30 minutes most days of the week; (2) eat 5 or 6 servings of fruits or vegetables daily; (3) smoke; (4) drink no more than 1 to 2 alcoholic drinks daily. Of the 500 respondents, 25% reported that they smoked; only 10% stated that they drank no more than 1 to 2 drinks daily. None of the respondents reported exercising at least 30 minutes per day or eating 5 or 6 servings of fruits or vegetables.

Ordinal responses are made to fit on a scale that is ordered from positive (*definitely or probably important*) to negative (*definitely or probably not important*). Ordinal data thus are often characterized by counts of the numbers and percentages of people who select each point on a graded scale.

> Of 500 respondents completing the question, 250 (50.0%) rated each preventive health behavior as definitely or probably important.

Field studies often ask for numerical data as when respondents are asked for their birth date. From the date, you can calculate each respondent's age. Age is

considered a numerical and continuous measure, starting with zero and ending with the age of the oldest person in the study. Numerical data lend themselves to many statistical operations. Typical findings might appear as follows: The average age of the respondents was 43 years. The oldest person was 79 years of age, and the youngest was 23.

10.3 How Are Field Study Measures and Questions Organized?

10.3.1 Length

The length of a measure depends upon what you need to know and how many questions you need to ask to get credible answers (Bourque and Fielder 2003a; Bourque and Fielder 2003b). Another consideration is the respondents. How much time do they have available, and will they pay attention? Relatively young children, for example, may only stay still for a few minutes, so shorter interviews, for example, may be better. You must also consider your resources: Longer measures may be more costly to design, validate and administer.

10.3.2 Question Order

All field measures should be preceded by an introduction, and the first set of questions should be related to the topic described in it. This is illustrated in Box 1.

Note that the interviewer starts off by saying that questions will be asked about satisfaction with the Health Clinic, and the first question calls for a rating of satisfaction.

In general, questions should proceed from easiest to answer and the most familiar to most difficult and least familiar. In a survey of needs for health services, items can first be asked about the respondent's own needs for services, then their family's, community's, and so on.

Questions of recall should also be organized according to their natural sequence. Do not ask very general questions: "When did you first start feeling dizzy?" Instead, prompt the respondent and ask: "In the past three months, how often did you felt dizzy? Think about the last time you felt dizzy. Was it in the morning, afternoon or evening?"

Sometimes the answer to one question will affect the content of another. When this happens, the value of the measure may be diminished (Example 3).

Box 1. An Introduction to a Telephone Interview and Its First Question

Hello. I am calling from the Health Clinic. We are surveying people who use the Health Clinic to find out whether it provides satisfactory services. Your name was selected at random from the Clinic's database. Our questionnaire will take no more than four minutes. You can interrupt me at any time. May I ask you the questions? [IF YES, CONTINUE. IF NO, SAY: Thank you AND HANG UP.]

CONTINUE HERE:

The first question asks you about your overall satisfaction with the Health Clinic. Do you consider it [READ CHOICES]

a. Definitely satisfactory
b. Probably satisfactory
c. Probably not satisfactory
d. Definitely not satisfactory

[DO NOT SAY]

e. No opinion or don't know/wrong answer

Example 3. *Ordering Questions*

Which question should come first?

a. How efficient is the nursing staff?

Or

b. Which improvements in nursing do you recommend?

Answer: Question b should come before Question a. If it does not, then the respondent might offer suggestions for the improvement of the nursing staff's efficiency of the nursing merely because it has been suggested. ♦

Place relatively easy-to-answer questions at the end of the measure. When questionnaires, for instance, are long or difficult, respondents may get tired and answer the last questions carelessly or not answer them at all. You can place demographic questions (age, income, gender, and other background characteristics) at the end because these can be answered quickly.

Avoid many items that look alike. Twenty items, all of which ask the respondent to agree or disagree with statements, may lead to fatigue or boredom, and the respondent may give up. To minimize loss of interest, group the questions and provide transitions that describe the format or topic. For instance, say or print something like: "The next set of questions ask about your use of health services."

Questions that are relatively sensitive should be placed toward the end. Topics such as grooming habits, religious views, and positions on controversial subjects such as abortion and assisted suicide must be placed far enough along so there is

reason to believe the respondent is willing to pay attention, but not so far that he or she is too fatigued to answer properly.

Finally, questions should appear to reasonable people to be in a logical order. Do not switch from one topic to another unless you provide a transitional statement to help the respondent make sense of the order.

Here is a checklist of points to consider in selecting the order for the questions in your field measure:

Checklist to Guide Question-Order

- For any given topic, ask relatively objective questions before the subjective ones.
- Move from the most familiar topics to the least.
- Follow the natural sequence of time.
- See to it that all questions are independent.
- Avoid many items that look alike.
- Ask sensitive questions well after the beginning.
- Place questions in a logical order.

10.3.3 Aesthetics and Other Concerns

A measure's appearance is important. A self-administered questionnaire that is hard to read can confuse or irritate respondents who may not answer accurately or at all, reducing the reliability of the responses. A poorly designed interview form with inadequate space for recording answers will reduce the efficiency of even the most skilled interviewers.

10.3.4 Branching Questions or Skip Patterns

What happens when you are concerned with getting answers to questions that you know are only appropriate for part of your group? Suppose you were interviewing older adults in a general practice to learn about their medication-use. You know that many of these patients are likely to be taking certain kinds of mediations such as antihypertensives, NSAIDs, and aspirin. However, some people will be taking all of these medications each day, while others will be taking none.

If you want to ask about a topic that you know in advance is *not* relevant to everyone in the study, you might design a form such as the one in Example 4.

Example 4. *Skip Patterns or Branching Questions*

- Do you take any of the following medications (a list is provided)
  a. *No* (Go to question 4)
  b. *Yes*

[If yes] How often do you take your usual dose (choices are given such as once a day; only when needed)?

OR

- Do you take any of the following medications?
  a. *Yes* (Complete Section A)
  b. *No* (Go to Section B) ♦

Skip patterns may be confusing to people and should be avoided in self-administered printed questionnaires. Interviewers must be trained to follow the branches. Computer-assisted and online questionnaires are effective vehicles for branching because you can design the software to guide the respondent. For instance, if the questionnaire tells the respondent, "If no, go to question 6," the respondent who answers "no" will automatically be sent to question 6.

What Does It Take to Ensure Proper Administration of Field Instruments? 10.4

Self-Administered Questionnaires 10.4.1

Self-administered questionnaires take the form of written, computer-assisted, and online surveys. They require a great deal of advance preparation and subsequent monitoring to get a reasonable response rate. These questionnaires are given or sent directly to people for completion. Advance preparation, in the form of careful editing and tryouts, is necessary in helping to produce a clear, readable self-administered questionnaire (Bourque and Fielder 2003a). You should always review the returns. Are you getting the response rate you expected? Are all questions being answered? The following is a checklist for using self-administered questionnaires:

Checklist for Using Self-Administered Questionnaires

- Mail respondents a letter or email them in advance telling them the purpose of your study. The letter should inform people to expect a questionnaire, explain the importance of the study and the respondent's role, list study supporters and sources of funding, and describe procedures to ensure confidentiality.
- Prepare a short, formal letter to accompany the questionnaire form. If you have already sent an advance letter, this one should be very concise.
- Offer to send respondents a summary of the findings so they can see just how the data are used. (If you promise this, allocate resources for it.)
- If you ask questions that may be construed as personal – such as gender, age, or income – explain why the questions are necessary.

- Keep the questionnaire procedures simple. Provide stamped self-addressed envelopes for written, mailed questionnaires. Make sure no special software is needed for online surveys (e.g., to download graphics). If special software is necessary, set up a system for ensuring that all respondents who are eligible for the survey have access to the software.
- Keep questionnaires as short as you can. Ask only questions you are sure you need and do not crowd them together. Give respondents enough room to write or check boxes and be sure each question is set apart from the next.
- Consider incentives. This may encourage people to respond. Incentives may range from certificates of appreciation and money and stamps to pens, fuel and food.
- Be prepared to follow up or send reminders. These should be brief and to the point. For mailed and online surveys, it often helps to send another copy of the questionnaire to non-respondents. Do not forget to budget money and time for these additional mailings.

10.4.2 Interviews

Finding Interviewers. Interviewers should fit in as well as possible with respondents. They should avoid flamboyant clothes, haircuts, and so on. Sometimes it is a good idea to select interviewers who are similar to respondents in gender, age or other demographic characteristics.

It is also important that the interviewers be able to speak clearly and understandably. Unusual speech patterns or accents may provoke unnecessarily favorable or unfavorable reactions. The interviewer's way of talking is of course an extremely important consideration in the telephone interview. The interviewer's attitude toward the study and the respondent will influence the results. If the interviewer does not expect much from the interview and sends this message, the response rate and reliability of responses will probably suffer. To make sure you are getting the most accurate data possible, you should systematically and frequently monitor the interviewers' progress (Bourque and Fielder 2003b).

Training Interviewers. The key to a good telephone or face-to-face interview is training (Frey 1989). The overall goal of training should be to produce interviewers who know what is expected of them and how to answer questions and also know where to turn if problems arise unexpectedly in the field.

Whether you are training two interviewers or twenty, it is important to find a time to meet together. The advantage of meetings is that everyone can develop a standard vocabulary and share problems encountered in the field.

Once at the training site, trainees must have enough space to sit and write or perform any other activities you will require of them. If you want them to interview one another as practice for their real task, be sure the room is large enough so that two or more groups can speak without disturbing the others. You may even need several rooms.

Trainees should be taken step by step through their tasks and given an opportunity to ask questions. It is also essential to tell them some of the reasons for their tasks so they can anticipate problems and be prepared to solve them. The most efficient way to make sure the trainees have all the information they need to perform their job is to prepare a manual. Here you can explain what they are to do and when, where, why, and how they are to do it.

Conducting Interviews. The following are suggested guidelines for conducting interviews.

- Make a brief introductory statement that will describe who is conducting the interview ("Dr. Mary Doe for Armstrong Memorial Medical Center"), tell why the interview is being conducted ("to find out how satisfied you are with our after-surgery program"), explain why the respondent is being called ("We're asking a random sample of people who were discharged from the hospital in the last two months"), and indicate whether or not answers will be kept confidential ("Your name will not be used without your written permission").
- Try to impress the person being interviewed with the importance of the interview and of the answers. People are more likely to cooperate if they appreciate the importance of the subject matter. Do not try to deal with every complaint or criticism, but suggest that all answers will receive equal attention.
- Check the hearing and "literacy" of the respondent. Although it is important to stay on schedule and ask all the questions, a few people may have trouble hearing and understanding some of the questions. If that happens, reappraise the eligibility of the respondent (perhaps an interview is not the best method of obtaining reliable data from this respondent; other methods may be more appropriate). Another option is to speak more clearly and slowly.
- Ask questions as they appear in the interview schedule. It is important to ask everyone the same questions in the same way or the results will not be comparable.

Monitoring Interview Quality. To make sure you are getting the most accurate data possible, you should monitor the quality of the interviews. This might mean something as informal as having the interviewer call you once a week, or something as formal as having them submit to you a standardized checklist of activities they perform each day. If possible, you may actually want to go with an interviewer (if it is a face-to-face interview) or spend time with telephone interviewers to make sure that what they are doing is appropriate for the study's purposes. To prevent problems, you might want to take some or all of the following steps:

Tips for Ensuring Quality

- Establish a hot line. This means having someone available to answer any questions that might occur immediately, even at the time of an interview. Consider obtaining a toll-free number.

- Provide written scripts. If interviewers are to introduce themselves or the study, give them a script or set of topics to cover. The script may have to be approved by an Institutional Review Board.
- Make sure you give out extra copies of all supplementary materials. If data collectors are to mail completed interviews back to you, for example, make sure to give them extra forms and envelopes.
- Provide an easy-to-read handout describing the purpose of the interview and the content of the questions.
- Provide a schedule and calendar so that interviewers can keep track of their progress.
- Consider providing the interviewer with visual aids. Visual aids may be extremely important when interviewing people in-person whose ability to speak or read may be limited. The preparation of audiovisual aids for use in an interview is relatively expensive and requires that the interviewers be specially trained in using them.
- Consider the possibility that some interviewers may need to be retrained and make plans to do so.

Computer-Assisted Telephone Interviews or CATI. Computer-assisted interviewing is becoming increasingly accepted as a useful field work tool. With CATI, the interviewer reads instructions and questions to the respondent directly from the computer monitor and enters the responses directly into the computer (Bourque and Fielder 2003b). The computer, not the interviewer, controls the progression of the interview questions. No paper copies of the interview are produced, thus eliminating the need to find secure storage place for completed questionnaires.

CATI software programs enable the researcher to enter all telephone numbers and call schedules into the computer. When the interviewer logs on, he or she will be prompted with a list of phone numbers to call, including new scheduled interviews and callbacks. For example, suppose the interviewer calls someone at 8 AM, but receives no answer. The CATI program can automatically reschedule the call for some other time. CATI programs are also available that enable specially trained interviewers to contact respondents with unique needs. For instance, suppose your study sample consists of people who speak different languages. CATI will allow multi-lingual interviewers to log on with certain keywords; the computer will then direct them to their unique set of respondents.

The major advantage of CATI is that once the data are collected they are immediately available for analysis. However, having easy access to data may not always be a blessing. Some researchers may be tempted to analyze the data before the completion of data collection, and the preliminary results may be misleading. A main value of easy access to data, certainly in the early stages of data collection, lies in having the means to check on the characteristics of the respondents and to monitor the quality of the CATI interviewers in obtaining complete data.

Intensive interviewer training is crucial when using CATI in field studies. Interviewers must first learn how to use the CATI software and handle computer problems should they arise during the course of the interview. For instance, what

will the interviewer do if the computer "freezes"? Further, the interviewer needs to practice answering questions invariably posed by respondents regarding the study's objectives, methods, human subjects' protections and incentives. In fact, given the complexity of CATI, training may take up to a week. Thus, at the present time, use of CATI should probably be considered primarily when a study is well funded because it is a relatively expensive and specialized form of data collection.

CATI takes two forms. The first, which is most commonly used, consists of a lab or a facility furnished with banks of telephone calling stations that are equipped with computers linked to a central server. The costs of building the lab are extremely high and include assembling soundproof cubicles and having either a master computer that stores the data from the individual computers or linkage to a server. Additional resources are needed to cover the cost of leasing CATI software and hiring a programmer to install it. Training for this type of CATI is expensive, requiring a great deal of practice. There are also numerous incidental costs including those for headsets, seats and desks, instructional manuals and service contracts for the hardware.

A second type of computer-assisted telephone interviewing system consists of CATI software programs that are run on laptops. With this type of CATI, the researcher only needs a laptop and access to a telephone connected to the Internet. In time, we can expect that the telephone will be superseded by wireless access to the Internet, making this type of telephone interviewing a desirable method for collecting data in the field and sending them to a central server. Moreover, this second type of CATI is appropriate for studies with a variety of funding levels because it is portable and relatively inexpensive. The portability of laptops, however, raises concerns about patient privacy. Laptops are sometimes shared or stolen, providing easy access to confidential respondent data. In anticipation of these concerns, laptops that are used for CATI should be dedicated to a single study, strict privacy safeguards must be enforced, and interviewers must receive special training to ensure proper CATI implementation and respondent protection. In the U.S., patient privacy rules have become increasingly strict (e.g., through the Health Insurance Portability and Accountability Act or HIPAA), with costly penalties for violation.

How Can You Assure a Reliable and Valid Field Measure? 10.5

Pilot Testing 10.5.1

Once a field measure has been assembled, it should be tested to determine the ease with which it can be administered and to estimate the accuracy of the data. Pilot testing includes evaluating the logistics of administration as well as the ease of use of the form itself. The purpose of the pilot test is to answer these questions:

Questions Answered by Pilot Testing

- Will the measure provide the needed information? Are certain words or questions redundant or misleading?
- Are the questions appropriate for the respondents?
- Will information collectors be able to use the forms properly? Can they administer, collect, and report information using any written directions or special coding forms?
- Are the procedures standardized? Can everyone collect information in the same way?
- How consistent or reliable is the information?

10.5.2 Reliability and Validity: The Quality of the Measure

A ruler is considered to be a reliable instrument if it yields the same results every time it is used to measure the same object, assuming the object itself has not changed. A yardstick showing that you are 6 feet 1 inch tall today and 6 feet 1 inch six months from today is reliable.

People change over time. You may be more tired, angry, and tense today than you were yesterday. People also change because of their experiences or because they learned something new, but meaningful changes are not subject to random fluctuations. A reliable instrument will provide a consistent measure of important characteristics despite background fluctuations. It reflects the "true" score – one that is free from random errors.

A ruler is considered to be a valid instrument if it provides an accurate measure (free from error) of a person's height. But even if the ruler says you are 6 feet 1 inch tall today and six months from now (meaning it is reliable), it may be incorrect, that is, invalid. This would occur if the ruler were not calibrated accurately, and you are really 5 feet 6 inches tall.

If you develop an instrument that consists of nothing more than asking a hospital administrator how many beds are in a given ward, and you get the same answer on at least two occasions, you would have an instrument that is reliable. But if you claim that the same instrument reflects the quality of medical care, you have a reliable measure of questionable validity. A valid measure is always a reliable one, but a reliable one is not always valid (Bernard 2000; Dawson and Trapp 2001).

10.5.3 Ensuring Quality: Selecting Ready-to-Use Measures

One way to make sure that you have a reliable and valid measure is to use one that someone else has prepared and demonstrated to be reliable and valid through careful testing. This is particularly important to remember if you want to survey attitudes, emotions, health status, quality of life, and health beliefs (Stewart and Ware 1992; McDowell and Newell 1996). These factors, and others like them, are

elusive and difficult to measure. To produce a truly satisfactory measure of health, quality of life, and human emotions and preferences thus requires a large-scale and truly scientific experimental study.

Reliability

10.5.4

In reviewing a published field instrument (also, in assessing the quality of a home-made form) you should ask the following questions about four types of reliability: test-retest, equivalence, internal consistency, and interobserver reliability.

Test-Retest Reliability. Does the instrument have test-retest reliability? One way to estimate reliability is to determine if someone taking the measure answers about the same on more than one occasion. Test-retest reliability is computed by administering the measure to the same group on two different occasions and then correlating the scores from one time to the next to obtain a correlation coefficient (*r* value). Usually, to be considered reliable, an instrument should obtain a correlation coefficient of at least 0.70 (Stewart and Ware 1992).

You can calculate test-retest reliability for single questions, subsets, or entire measures. For example, suppose you are studying the use of medications in a sample of older adults. The instrument you are using asks this question, "How many prescription medications do you usually take each day?" In order to assess the consistency of the respondents' answers, you would ask the same question twice: at baseline and then a second time, say 2–4 weeks later. Assuming medication-use rates in your sample tend to be stable over short periods of time, any differences in responses to the question can be assumed to reflect measurement error and not changes in the use of medications. To calculate test-retest reliability for an entire measure, you would administer its entire set of questions at two different points in time, score it, and then calculate the correlation coefficient for the two scores.

Equivalence. Are alternative forms equivalent? If two different forms of a questionnaire are supposed to measure the same attitude, for example, you should make sure that people are likely to obtain the same score regardless of which one they take. If you want to use Form A of the instrument as a premeasure, for example, and Form B as a postmeasure, check the equivalence of the two forms to make sure one is not different from the other.

Equivalence reliability can be computed by giving different forms of the instrument to two or more groups that have been randomly selected. The forms are created either by using differently worded questions to measure the same attributes or by reordering the questions. To test for equivalence, you can administer the different forms (reordered or reworded) at separate time points to the same population, or if the sample is large enough, you can divide it in half and administer each of the two alternate forms to half of the group. In either case, you would first compute mean scores and standard deviations on each of the forms and then correlate the two sets of scores to obtain estimates of equivalence. Equivalence reliability coefficients should be at least 0.70.

Internal Consistency. Another measure of reliability is how internally consistent the questions are in measuring the characteristics, attitudes, or qualities that they are supposed to measure. To test for internal consistency, you calculate a statistic called coefficient alpha, or Cronbach's alpha, named for the person who first reported the statistic.(Anastasi 1982; Bernard 2000). Coefficient alpha describes how well different items complement each other in their measurement of the same quality or dimension.

Many researchers are not at all concerned with internal consistency because they are not going to be using several items to measure one attitude or characteristic. Instead, they are interested in the responses to each item. Decide if your instrument needs to consider internal consistency.

Example 5. *Internal consistency*
Internal consistency is important

A ten-item interview is conducted to find out patients' satisfaction with medical care in hospitals. High scores mean much satisfaction; low scores mean little satisfaction. To what extent do the ten items each measure the same dimension of satisfaction with hospital care?

Internal consistency is not important

A ten-item interview is conducted with patients as part of a study to find out how hospitals can improve. Eight items ask about potential changes in different services such as the type of food that might be served, the availability of doctors, nurses, or other health professionals, and so on. One item asks patients for their age, and one asks about education. Since this interview is concerned with views on improving eight very different services and with providing data on age and education of respondents, each item is independent of the others. ♦

Interobserver Reliability: Kappa. Kappa is a statistic used to measure interrater (or intrarater) agreement for nominal measures (Cohen 1960). Suppose two researchers are asked to independently review a sample of 100 medical records to determine health services utilization among a sample of diabetic patients. Suppose also that the reviewers are required to use a standardized form containing questions about utilization. One of the questions asks the reviewer to indicate whether or not each patient has visited the emergency department (ED) within the past month. Here are the reviewers' responses to that question (Table 10.1).

This is shown in the following formula in which O is the observed agreement and C is the chance agreement.

Measuring Agreement Between Two Coders: The *Kappa* (κ) Statistic

$$\kappa = \frac{O - C(\text{Agreement beyond chance})}{1 - C(\text{Agreement possible beyond chance})}.$$

Table 10.1. Reviewers' response

	Reviewer 1		
Reviewer 2	**No, Did not Visit ED**	**Yes, Did Visit ED**	
No	20^{C}	15	35^{B}
Yes	10	55^{D}	65
	30^{A}	70	

Reviewer 1 says that 30 (^{A}) of the 100 patients did not visit the ED, while Reviewer 2 says that 35 (^{B}) did not. The two reviewers agree that 20 (^{C}) patients did not visit the ED.
What is the best way to describe the extent of agreement between the reviewers? 20 of 100 or 20% (^{C}) is probably too low because the reviewers also agree that 55% (^{D}) of patients did visit the ED. The total agreement: 55% + 20% is an overestimate because with only two categories (yes and no), some agreement may occur by chance.

Here is how the formula works with the above example.

1. Calculate how many records the reviewers may agree by chance indicate that patents did not visit the ED. This is done by multiplying the number of no's and dividing by 100 because there are 100 interviews: $30 \times 35/100 = 10.5$
2. Calculate how many interviews they may agree by chance indicate that patients do visit the ED. This is done by multiplying the number of yes's and dividing by 100 : $70 \times 65/100 = 40.5$
3. Add the two numbers obtained in steps 1 and 2 and divide by 100 to get a proportion for *chance agreement*: $(10.5 + 45.5)/100 = 0.56$.

The *observed agreement* is 20% + 55% = 75% or 0.75. Therefore the agreement beyond chance is $0.75 - 0.56 = 0.19$: the numerator.

The *agreement possible beyond chance* is 100% minus the chance agreement of 56% or $1 - 0.56 = 0.44$: the denominator

$$\kappa = \frac{0.19}{0.44} = 0.43 \ .$$

What is a "high" *kappa*? Some experts have attached the following qualitative terms to *kappas*: 0.0–0.2 = slight; 0.2–0.4 = fair; 0.4–0.6 = moderate; 0.6–0.8 = substantial, and 0.8–0.10 = almost perfect.

Here are some questions to ask about a published field instrument's validity:

Questions to Ask About a Published Instrument's Validity 10.5.5

1. Does the instrument have predictive validity? You can validate an instrument by proving that it predicts an individual's ability to perform a given task or behave in a certain way. For example, a medical school entrance examination has predictive validity if it accurately forecasts performance in medical school. One way of establishing predictive validity is to administer

the instrument to all students who want to enter medical school and compare these scores with their performance in school. If the two sets of scores show a high positive or negative correlation, the instrument has predictive validity.

2. Does the instrument have concurrent validity? You can validate an instrument by comparing it against a known and accepted measure. To establish the concurrent validity of a new measure of quality of care, you could administer the new instrument and an already established, validated instrument to the same group and compare the scores from both instruments. You can also administer just the new instrument to the respondents and compare their scores on it to experts' judgment of the respondents' attitudes. A high correlation between the new instrument and the criterion measure (the established instrument or expert judgment) means concurrent validity. A concurrent validity study is only valuable if the criterion measure is convincing.
3. Does the instrument have content validity? An instrument can be validated by proving that its questions accurately represent the characteristics or attitudes that they are intended to measure. An instrument that is designed to measure health beliefs has content validity, for example, if it contains a reasonable sample of facts, words, ideas, and theories commonly used when discussing or reading about the formation of beliefs about disease or health. Content validity is usually established by consulting the literature and by asking experts and prospective respondents whether the questions represent the knowledge, attitudes and behaviors you want to measure.
4. Does the instrument have construct validity? Construct validity means that the instrument measures what it purports to and not something else. Because of the difficulty of obtaining a true measure of the concepts and ideas that characterize epidemiological studies, construct validity must be established experimentally. One method of doing this is to administer the instrument to people whom selected experts say exhibit the behavior associated with the construct. Usually, the experts based their judgments on theories that have empirical support and on clinical experience. If the people chosen by the experts to be exemplars of the behavior also obtain a high score (i.e., a higher score means greater evidence of the behavior), then the instrument is considered to have construct validity. This form of validity is usually established after years of experimentation and experience with the measure.

10.5.6 Suggested Guidelines For Pilot Testing

- Try to anticipate the actual circumstances in which the instrument will be conducted and make plans to handle them. For interviews, this means reproducing the training manual and all forms; for online surveys and mailed questionnaires, you have to produce any cover letters, return envelopes, and so on. Needless to say, this requires planning and time and can be costly.

- You can start by trying out selected portions of the instrument in a very informal fashion. Just the directions on a self-administered questionnaire might be tested first, for example, or the wording of several questions in an interview might be tested. This is sometimes called a "cognitive" pretest.
- Choose respondents for the pilot who are similar to the ones who will eventually complete the measure. They should be approximately the same age, with similar education, and so on.
- Enlist as many people in the trial as seems reasonable without wasting your resources. Probably fewer people will be needed to test a five-item questionnaire than a twenty-item one.
- For reliability, focus on the clarity of the questions and the general format of the instrument. Look for:
  - failure to answer questions
  - giving several answers to the same question
  - writing comments in the margins.

Any one of these is a signal that the measure may be unreliable and needs revision. Are the choices in forced-choice questions mutually exclusive? Have you provided all possible alternatives? Is the questionnaire or interview language clear and unbiased? Do the directions and transitions make sense? Have you chosen the proper order for the questions? Is the questionnaire too long or hard to read? Does the interview take too long? For instance, you planned for a ten-minute interview, but your pilot version takes twenty.

Consider this: In a pilot of a self-administered survey of children's health behaviors, respondents were asked how often they washed their hands after eating. All six children between 8 and 10 years of age answered "always" after being given the choices "always," "never," and "I don't know." The choices were changed to "almost always," "usually," and "almost never." With the new categories, the same six children changed their answers to two "almost always" and four "usually."

Field Work Language and Culture

10.6

If you plan on translating an existing data collection instrument or measure, do not assume that you can automatically reword each question into the new language. Between the original language and the next language often lie cultural gaps. You may need to reword each question.

To avoid confusing people and even insulting them because you misunderstand their language or culture, you should follow a few simple guidelines. These involve enlisting the assistance of people who are fluent in the language (and its dialects) and pilot testing the measure with typical respondents. Suggested guidelines for translation always include the following.

10.6.1 Suggested Guidelines for Translation

- Use fluent speakers to do the first translation. If you can, use native speakers. If you can afford it, find a professional translator. The art of translation is in the subtleties – words and phrases that take years and cultural immersion to learn. If you use fluent speakers, you will minimize the time needed to revise question wording and response choices.
- Tryout the translated measure with 3 to 5 native speakers. Ask: What is this question asking you? Can you think of a better way to ask this question?
- Revise the measure with the help of the original translator.
- Translate the measure back into the original language. Use a different translator for this task. Does this "back translated" instrument match the original version? If not, the two translators should work together to make them match.
- Try the resulting measure on a small group (5–10) of target respondents. If the two translators could not agree on wording, let the group decide.
- Revise the measure.
- Pilot test the measure.
- Produce a final version.

10.7 Managing the Data

Data management consists of the methods used to store and organize information so that it can be analyzed. Data management starts with an analysis plan and ends with the final analytic operation. The analysis plan describes the hypotheses to be tested or research questions that will be answered. The plan is a guide to the data that will be collected, entered and subsequently analyzed (Example 6).

Example 6. *A Portion of an Analysis Plan for an Interview on Health and Alcohol Use in the Elderly*

Hypothesis: More men than women will exceed drinking limits.

Variables: gender; drinking limits

Planned Analysis: Chi square to test for differences between numbers of men and women who exceed limits ♦

Modifications to the original analysis plan can be expected, especially in large studies with a great deal of data.

A second data management activity is the creation of a code book. The contents of a code book may vary among researchers. Some researchers include in their definition only a description of study's variables (such as [DRINK]) and how they are categorized or labeled (such as 1 = 1 to 2 drinks daily; 2 = 3 or more drinks

daily; 9 = no data). Increasingly, many investigators are promulgating the view that code books should include all the information needed to reproduce the study. For example (Example 7), the Field Survey, a large California polling survey group posts information like the following on its web site (http://field.com/fieldpoll/).

Example 7. *Table of Contents for A Code Book*

I. Methods
   A. *Sampling*
      1. Sampling design to include eligibility criteria (e.g., 65 years of age and older; have had at least one drink in the past month)
      2. Sampling strategies (e.g., stratified random sampling; convenience or opportunistic sampling; etc.)
      3. Sample size and justification
      4. Recruitment and enrollment
      5. Sampling statistics to include weight and sampling error calculations
   B. *Human subjects*: Informed consent
   C. *Research design* or how participants were assigned to groups (if appropriate); number and timing of instrument administration

II. Data Collection
   A. A copy of the instrument
      6. The origins of the questions (e.g., adapted from a published instrumement; created for this one)
      7. Description of how each response is coded (e.g., 1 = yes; 2 = no; 9 = no data)
   B. Training of data collectors; quality control
   C. Information on reliability and validity

III. A Data File Description
   The variable names [DRINK], Labels (quantity and frequency) and values and value labels (1 = 1 to 2 drinks daily; 2 = 3 or more drinks daily; 9 = no data) ♦

A major problem in data management is how to handle missing data. Say, you mail 100 questionnaires and get 95 back. Is this a 95% response rate? Suppose that upon close examination, you discover that half the respondents did not answer question 5, and that none of the questions was answered by all respondents. With all that missing information, you cannot claim to have a 95% response rate.

What should be done about missing responses? In some situations, you may be able to go back to the respondents and ask them to answer the questions they omitted. In small studies, where the respondents are known, the respondents may be easily re-contacted. But collecting information a second time is usually impractical, if not impossible, in most studies. Some studies are anonymous,

and you do not even know who the respondents are. In institutional settings, you may have to go back to the Institutional Review Board to get permission to contact the respondents a second time. This may take too much time for your purposes.

Computerized questionnaires can be programmed so that the respondent must answer one question before proceeding to the next. Some respondents may find this approach frustrating, however, and refuse to complete the questionnaire. Although compelling respondents to answer all questions is touted as a major advantage of computer-assisted data collection, some researchers believe that forcing respondents to answer every question is coercive and unethical.

A key management activity is data entry, that is, the process of getting data into the computer. It usually takes three forms. In the first, someone enters data from a coded instrument into a database management program or spreadsheet. The data are then saved in as text or ASCII files, so that they can be exported into a statistical program like SPSS, SAS, or Stata. A second type of data entry involves entering data directly into a statistical program like SPSS, SAS, or Stata. In the third form of data entry, the respondent or interviewer enters responses directly into the computer. Data entry of this type is associated with computer-based measures including CATI and online surveys. The responses are automatically entered into database management systems or statistical programs (usually through special translation software). Programs are also available that will automatically convert one file format into another (say from SAS to Stata).

Database management programs, statistical programs and computer-assisted data collection with automatic data entry can facilitate accuracy by being programmed to allow the entry of only legal codes. For instance, if the codes should be entered as 001–010, then you can write rules so that an entry of 01 or 10 is not permitted. If you try to enter 01 or 10, you will get an error message. With minimum programming, the program can also check each entry to ensure that it is consistent with previously entered data and that skip patterns are respected. That is, the program can make sure that the fields for questions that are to be skipped by some respondents are coded as skips and not as missing data. Designing a computer-assisted protocol requires skill and time. No protocol should be regarded as error-free until it has been tested and retested in the field.

Once the data are entered, they need to be cleaned. A clean data set can be used by anyone to get the same results as you do when you run the analysis. Data become "dirty" for a number of reasons including miscoding, incorrect data entry and missing responses.

To avoid dirty data, make sure that coders or data entrers are experienced, well trained, and supervised. Check variable values against preset maximum and minimum levels, so that if someone enters 50 instead of 5, the maximum, you know there is an error. You can also minimize errors by making sure your coding scheme distinguishes among truly missing (no response or no data), from don't know and not applicable.

Run frequencies on your data as soon as you have about 10% of the responses in. Run them again and again until you are sure that the fieldwork is running smoothly. Frequencies are tabulations of the responses to each question. If your data set is relatively small, you can visually scan the frequencies for errors. For large databases with many records, variables, skip patterns, and open-ended text responses, you may need to do a systematic computerized check. All leading statistical programs provide for cleaning specifications that can be used during data entry and later as a separate cleaning process.

Several other problems may require you to clean up the data. These include having to deal with the complete absence of data because some questionnaires have not been returned, for instance, with missing data from questionnaire that have been returned, and with questionnaires that contain data that are very different form the average respondents.

What Are Reasonable Resources? 10.8

Fieldwork resources are reasonable if they adequately cover the financial costs of and time needed to conduct all activities in the time planned. This includes the costs of, and time for, hiring and training staff, preparing and validating forms, administering the instrument or measure, and analyzing, interpreting, and reporting the findings.

How much does it cost to conduct fieldwork? This question can be answered by obtaining the answers to seven other questions.

1. What are the major tasks?
2. What skills are needed to complete each task?
3. How much time do I have?
4. How much time does each task take?
5. Whom can I hire to perform each task?
6. What are the costs of each task?
7. What additional resources are needed?

Here is a checklist of typical field work tasks:

Field Work Task Checklist

- *Prepare the instrument for use in the field.*
  - Identify existing and appropriate instruments.
  - Conduct a literature review.
  - Contact other researchers.
  - Adapt some or all questions from existing instruments.
  - Prepare a new instrument.

- *Identify, enroll and recruit subjects.*
  - Determine eligibility criteria (who should be included and excluded).
  - Determine sample size.

- Identify sources for identifying respondents (e.g., existing data bases; patients in a waiting room; patients with appointments).
- Devise plans for coordinating respondents' willingness to participate and the study's field work needs. For instance, you may have to provide transportation for interviewers or participants.

- *Prepare documents for the IRB (Institutional Review Board).*
  - Develop procedures for insuring ethical principles of research. Such principles include respect for people and their autonomy; protecting people from harm and taking active steps to protect them; and balance potential risks with benefit from participation in the study.
  - Devise methods for ensuring the protection of the people who participate in field studies, including the preparation of fliers, recruitment letters, and informed consent forms.
  - Make provisions for protection against research misconduct including exaggerating findings or releasing them without permission.

- *Pretest the Instrument.*
  - Identify a relatively small sample for the pretest.
  - Conduct a "cognitive" pretest by going over the instrument question by question with each respondent.

- *Pilot test the instrument.*
  - Identify the sample for the pilot test.
  - Obtain permission for the pilot test.
  - Analyze the pilot-test data.
  - Revise the instrument to make it final.

- *Administer the instrument.*
  - Hire staff.
  - Train staff.
  - Monitor the quality of administration.
  - Retrain staff.
  - Send out mail, supervise the questionnaire, conduct interview.
  - Follow up.

- *Manage the data.*
  - Code responses.
  - Prepare code book.
  - Consult with programmer.
  - Train data enterers.
  - Enter the data.
  - Run a preliminary analysis.
  - Clean the data.
  - Prepare a final codebook.

- *Analyze the data.*
  - Prepare an analysis plan.
  - Analyze the reliability and validity of the instrument.
  - Analyze the results of the study.

- *Report the results.*
  - Write the report.
  - Have the report reviewed.
  - Modify the report based on the reviews.
  - Prepare presentation.
  - Present the report orally.

Who Will Do It, and What Resources Are Needed? Personnel, Time, And Money

10.8.1

Fieldwork happens because one or more persons are responsible for completing the required tasks. In a very small study, one or two persons may be in charge of developing field instruments, administering them, analyzing the data, and reporting the results. In larger studies, teams of individuals with differing skills are involved. Sometimes, a study is planned and conducted by the staff with the assistance of consultants who are called in for advice or to complete very specific activities.

First, you need to plan the activities and tasks that need to be completed. Once this is accomplished, you then decide on the skills required for each task. Next, you decide on the specific personnel or job descriptions that are likely to get you as many of the skills you need as efficiently as possible. For example, suppose your study design requires someone with experience in training interviewers and writing questions. You may just happen to know someone who needs a job and has both skills, but if you do not know the right person, knowing the skills needed will help you target your employment search.

The specific resources that are needed for each study will vary according to its size and scope and the number of skills and personnel needed to accomplish each task. Example 8 illustrates the types of skills and resources for a "typical" field study.

Example 8. *Tasks, Skills, and Resources: An Explanation*

1. *Prepare the instrument for use in the field.*
   If an instrument is to be adapted from an already existing instrument, expertise is needed in conducting literature reviews to find out if any potentially useful instruments are available. Sometimes, a reasonably good instrument is available: Why spend time and money to prepare an instrument if a valid one exists? It helps to have experience in the subject matter being addressed and to know who is working in the field and might either have instruments or questions or know where and how to get them.

Selecting items or rewording them to fit into a new measure requires special skills. You must be knowledgeable regarding the respondents' reading levels and motivation and have experience writing questions.

Preparing an entirely new instrument is daunting. A job description for an instrument writer would call for excellent writing skills and knowledge of all topics being assessed.

2. *Prepare Materials for the Institutional Review Board (IRB).*
U.S. researchers cannot perform research with funds from the U.S. government without approval from an Institutional Review Board or IRB – an independent group of people whose job is to evaluate if proposed research conforms to ethical principles (Brett and Grodin 1991). The IRB typically require a written explanation of the study plans (including rationale, purposes, and methods); the field forms and one or more informed consent forms. The following informed consent form has been approved by an IRB. As you can see, it provides potential respondents with descriptions of the study's purposes, the nature and characteristics of the tasks that will be required and describes procedures for ensuring confidentiality.

Box 2. Sample Consent Form

The Prostate Cancer Network (PROCANE)

You are asked to take part in three telephone interviews and three self-administered questionnaires on your general health, your quality of life since being diagnosed with prostate cancer, and the quality of healthcare you have received while in the PROCANE Program. XXX MD, MPH is directing the PROCANE research study. Dr. XXX works in the Department of Urology at the University of YYYY. You are being asked to take part in the interviews and questionnaires because you are enrolled in the PROCANE program. You can choose to take part in this study or not. If you volunteer to take part in this study, you may stop taking part in the study at any time. This will have no effect of any kind on the health care you receive through the PROCANE program.

Disclosure Statement

Your health care provider may be an investigator in this research protocol. As an investigator he/she is interested in both your clinical welfare and your responses to the interview questions. Before entering this study or at any time during the study, you may ask for a second opinion about your care from another doctor who is in no way associated with the PROCANE program. You are not under any agreement to take part in any research project offered by your physician.

Reason for the Telephone Interviews and Self-Administered Questionnaires

The interviews and the questionnaires are being done for the following reason: To find out if the PROCANE program is meeting the needs of the patients enrolled in the program. During the telephone interview, a trained member of the PROCANE staff will ask you a series of questions about:

- Your health
- Your quality of life since being diagnosed with prostate cancer, and
- The quality of the healthcare you have received while in the PROCANE program.

The self-administered questions will cover the same topics. But, you will be able to answer them on your own.

The PROCANE program will use your answers and the answers from other program participants to find out if the program is providing the right services to its participants and to find out if any changes need to be made to the program.

What You Will Be Asked to Do If you agree to take part in this study, you will be asked to do the following things:

1. Answer three short (20 minutes) telephone interviews. The telephone interviews will ask you general questions about your health, your quality of life since being diagnosed with prostate cancer, and the quality of healthcare you have received while in the PROCANE program. You will be called to complete an interview when you first enroll in PROCANE, 6 months after your enrollment, and when you leave the PROCANE program. The interviews will be completed at whatever time is best for you.

   Sample questions:
   - How confident are you in your ability to know what questions to ask a doctor?
   - During the PAST 4 WEEKS, how much did pain interfere with your normal work (including both work outside the home and housework)? Would you say not at all, a little bit, moderately, quite a bit, or extremely?
   - How much of the time during the LAST 4 WEEKS have you wished that you could change your mind about the kind of treatment you chose for prostate cancer?

2. Answer three short (20 minutes) self-administered questionnaires. The self-administered questionnaires will ask you general questions about your health, your quality of life since being diagnosed with prostate cancer, and the quality of healthcare you have received while in the PROCANE program. The self-administered questionnaires will be mailed to you when you first enroll in PROCANE, 6 months after your enrollment, and when you leave the PROCANE program. The self-administered questionnaires can be completed at whatever time is best for you. A pre-paid envelope will be provided to you in which to return each questionnaire.

   Sample questions:
   - Over the PAST 4 WEEKS, how often have you leaked urine?
   - Overall, how big a problem have your bowel habits been for you during the LAST 4 WEEKS?
   - Overall, how would you rate your ability to function sexually during the LAST 4 WEEKS?

3. If you do not understand a question or have a problem with a self-administered questionnaire, you will be asked to call Ms. AAA at the PROCANE office at **1-800-000-000.** She will be able to assist you.

Possible Risks and Discomforts

You may be sensitive about answering questions that ask about your physical and emotional health or your experiences with the PROCANE program. However, you do not have to answer any question with which you are uncomfortable.

Potential Benefits to Subjects and/or to Society

The purpose of the telephone interviews and self-administered questionnaires is to improve the services that PROCANE provides to the men enrolled in the program. Your responses might lead to changes in the program that would improve the services that PROCANE provides.

Payment for Taking Part

No payment will be given to you for completing the telephone interviews or self-administered questionnaires.

Confidentiality

Any information that is collected from you and that can be identified with you will remain confidential. Your identity will not be revealed to anyone outside the research team unless we have your permission or as required by law. You will not be identified in any reports or presentations. Confidentiality will be maintained in the following ways:

- All interviews and questionnaires will be coded with a number that identifies you. Your name will not be on any of these materials.
- A master list of names and code numbers will be kept in a completely separate, confidential, password-protected computer database.
- All copies of the self-administered questionnaires will be kept in a locked file cabinet in a locked research office.
- All telephone interviews will be recorded in a confidential computer database.
- When analysis of the data is conducted, your name will not be associated with your data in any way.
- Only research staff will have access to these files.

Taking Part and Choosing Not to Take Part in Telephone Interviews and Self-Administered Questionnaires

You can choose whether to take part in this study or not. If you decide to take part in the telephone interviews and self-administered questionnaires you may stop taking part at any time. This will have no effect of any kind on the health care you receive through the PROCANE program. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

Identification of Investigators

If you have concerns or questions about this study, please contact XXX, M.D., MPH, by mailing inquires to Box 000, Los Angeles, CA 900000-9990. He can be also reached at 1-800-000-000.

Rights of Participants

You may choose to end your agreement to take part in the telephone interviews and self-administered questionnaires at any time. You may stop taking part without penalty. You are not giving up any legal claims, rights or remedies because you take part in the telephone interviews and self-administered questionnaires. If you have questions about your rights as a research subject, contact the Office for Protection of Research Subjects, 2107 QQQ Building,, Box 951694, Los Angeles, CA 90095-1694, (310) 999-9999.

I understand the events described above. My questions have been answered to my satisfaction, and I agree to take part in this study. I have been given a copy of this form.

_____
Name of Subject (Please Print)

_____
Signature of Subject Date

Many IRB's also require detailed explanations of why each question on an instrument was chosen, how the study's participants were selected, and how you plan to ensure privacy for the respondent.

3. *Identify, recruit, and enroll patients*
   Participants in field studies can be identified from existing databases (e.g., Medicare data base; physician specialty membership lists; patient appointment logs), and they can be approached in the field (e.g., a clinic's waiting room). To obtain a valid sample, the research must establish eligibility criteria that describe who will be included and excluded into the study. Effective procedures must be devised to approach potential participants, screen for eligibility, inform eligible respondents about their role in the study, and enlist their cooperation. Often, these procedures need to be tested and retested until they achieve their desired goals. In the U.S., any contact with research subjects, including letters informing potential participants about a study or "scripts" to screen or enroll participants must be approved by an IRB.
   Recruitment letters are often sent in advance of a study to inform respondents of the study and its purposes. The following is an example of a recruitment letter for a study of staffing and clinical policies regarding labor and delivery in all hospitals in a single U.S. state. The letter was approved by an IRB and sent to the appropriate nurse manager at all hospitals in the state that delivered babies in a given year.

Box 3. Recruitment letter

<Letterhead and Logos>

Address
_____
_____
_____

Date _____

Dear Colleague (use name)

We are writing to invite you to participate in an exciting federally funded study of the range of clinical policies on Labor and Delivery (L&D) units in this state. Your

participation includes taking part in a structured interview regarding policies and procedures on your (L&D) unit.
As you are well aware, much interest has been placed on nurse staffing policies in general, but to date most studies have centered on medical and surgical units. This will be the first study that is specifically designed to identify staffing patterns used on Labor and Delivery units, and to associate these staffing patterns with clinical policies and patient outcomes.

Likewise, there have been numerous isolated studies regarding the role of nursing support in labor, the effectiveness of doulas, nurse midwives, and various techniques of labor management (e.g. active management, "walking epidurals," hydrotherapy, etc). To date, however, there has been no systematic attempt to describe what is really happening on L&D units in the "real world." At the completion of our study, we hope to be able to answer the following types of questions:

- What does "active management" mean to you?
- How prevalent is hydrotherapy?
- What types of clinicians are trained in "teaching" hospitals?
- When do physicians need to be "in house"?
- How are L&D units staffing in the current environment of a nursing shortage?
- What strategies are being used to monitor cesarean rates?
- How are staffing and clinical policies related to maternal and neonatal outcome?

If you agree to participate, we will send you an advance copy of the L&D Clinical Policy Survey, and call to arrange a convenient interview time for you. You will be compensated $60.00 for your participation, which will take about 45 minutes to one hour. Responses will be collated and serve as the first comprehensive overview of staffing and clinical policies on L&D units in this state. This project has been reviewed and approved by the Medical Center Institutional Review Board, and an information letter has been included for your review.

This project, has been funded by the Agency for Health, and has been widely endorsed by representatives of agencies promoting a better understanding of healthcare practices, healthcare quality, and healthcare outcomes including, but not limited to the following:

- **Mary Smith**, Administrative Director of the Association of this State's Nurse Leaders
- **Tom Rodriquez**, CEO Medical Center, Current Board of the Hospital Association, and Past President of the State Hospital Association
- **Robert Johnson MD, MPH**, Center for Disease Control Director Division of Birth Defects and Developmental Disabilities
- **William Roberts, MD**, College of Obstetricians and Gynecologists, Chair District XXX

If you have any questions regarding this project, please feel free to contact us at 310 666-789. We thank you in advance for your participation in this exciting project.

Sincerely,

Yvonne Bree, RN, DrPH
Vice President & Chief Nursing Officer
YYY Medical Center

Mathilde Grun, MD, MPH Director Maternal Fetal Medicine & Women's Health Services Research
YYYY Medical Center
Department Obstetrics & Gynecology

Recruitment letters tend to be most effective if they follow these suggestions:

1. Write the letter on letterhead. If possible place one or more logos on the stationery. The logos may be of the university or agency at which the study takes place and one or more of its supporters.
2. Personalize the letter, if appropriate. In the U.S., "Dear John," is often handwritten over the "Dear Dr. Jones" to express collegiality.
3. Describe the purpose of the study. In this case, examples of study questions are included.
4. Describe the role each participant will play. This letter informs the respondent that joining the study means participation in an interview that may last up to 60 minutes.
5. Describe the incentives you are prepared to offer the respondent. In this case, the incentive is financial.
6. Inform the respondent about confidentiality. According to the letter, the Medical Center's Institutional Review Board has approved the study. An information sheet is to be included with this letter describing how confidentiality is to be ensured. If in doubt, or you do not have an information sheet, the letter should include a statement about protection of privacy.
7. Describe the source of funding for the study.
8. Give the names of any agencies or organizations that endorse or co-sponsor the study.

Sometimes, recruitment is done by telephone. As seen below, recruitment also means collecting data on refusers. Such data are used to determine if the recruitment approach is effective and to provide information on the similarity between persons who agree to participate and those who do not. If differences exist between participants and refusers, the external validity of the study may be compromised.

Box 4. Parent and Child Telephone Recruitment Script

Hello, my name is [*FIRST AND LAST NAME*] and I am calling on behalf of the LAUSD/UCLA study, *Finding Solutions Together*. May I speak to [*NAME OF PARENT*]?

1. If someone other than respondent asks why you are calling, say:
   I'm calling about a research study being conducted by the school district.

   (Check one answer)
   a. *No one by that name is at this number* → Ask Q2
   b. *R not available* → Skip to Q3
   c. *I am speaking to the parent or the parent comes to the phone* → Skip to Q9
   d. *Refusal* → Skip to Q8

2. Confirm you have dialed correctly. Ask if the respondent was ever at this number and if they have a new number for the person you are trying to reach. If your informant cannot give you a new number, try directory assistance for a new listing. If no new number is listed, note as not located.

3. Say: Is there a more convenient time to reach R?
   a. *Yes* → Continue
   b. *No* → Go to Q5

4. Set call back appointment
   Date: _____
   Time: _____
   Say: Okay. We will try back at that time. Thank you. End call.

5. Say: Is this the best number to reach R or do you have a better number for him/her?
   a. *Yes* → Continue with Q7
   b. *No* → Go to Q6

6. Record new number: _____ → Say: Okay, we will try him/her at this number. Thank you for your help. End call.

7. Say: Okay, thank you. We will try again another time. May I leave you my name and toll free number in case R wants to call me back?
   a. *Yes* → Provide name and toll free number. End call.
   b. *No* → Thank the informant and end call.

8. Say: Thank you very much for your time. End call. Fill out information below:

   Refusal information

   Who did you speak to? _____

   Reason for refusal?

   _____

   Hello, my name is [*NAME*] *(if not already introduced)* and I am calling from [*NAME OF SCHOOL*]. I am calling you today because your child participated in the first part of our study and [*NAME OF CHILD*] reported on the questionnaire that [*HE/SHE*] has had difficulties related to stressful experiences that may benefit from our program. I am calling to see if you and your child are interested in participating in a program that can help children learn ways of coping with stressful experiences.

9. Say: Are you interested in hearing more about this study?
   a. *Yes* → Continue
   b. *No* → End the call → Skip to Q13

   We are conducting a study of youth in middle school who have experienced a very stressful event. The goal of the study is to find out ways that young people react to stressful life events, and whether a new program might help them feel better.
   If you and your child volunteer to be in this study, we would ask you to do the following things:
   - Your child would be given the opportunity to attend a group at school for children who have had stressful experiences and who could benefit from learning ways to improve the way that they feel and act. These groups will have 5–10 students and a group leader and will meet once a week at school for 10 weeks. The group sessions will be audio taped for research purposes

only. There will also be one meeting between your child and the counselor alone, about halfway through the program and four optional parent meetings for parents to learn more about how to help their child at home.
- Not all children who qualify for this program will be in the program right away. Children will be chosen at random, like a flip of a coin, to start right away in the program or to receive the program in about 3 months. Those children who do not get into the program right away will also be offered other services to help them while they wait to start this school program.
- In addition to the groups, we will ask that you and your child meet with an interviewer to answer questions about background about your child and family and how your child has been feeling recently. The parent interview will take about one hour and will be set up at a time and place that is convenient for you. (FILL IN DESCRIPTION OF QUESTIONS) The child interview will take about 30 minutes at school (FILL IN DESCRIP). We will ask you and your child to complete this interview before the program starts, and again at 3 months and at 6 months after starting the program, for a total of 3 times.
- We will ask your child's teacher to complete a short checklist about your child's behavior at school before the program, at 3 months, and at 6 months.

10. SAY: Are you willing to meet with me to discuss your child's participation in this study in more detail, and if interested, complete the first parent interview?
    a. *Yes* → GO TO Q11
    b. *No* → GO TO Q13

11. SAY: Would you prefer to meet at the school or at your home?
    CHECK ONE ANSWER:
    a. *School*
    b. *Home (if home, obtain address)*: _____
    _____
    _____

    SAY: What is a good day and time for [you and/or your child] to do the interview?

    Date: _____

    Time: _____

    IF NECESSARY, OTHERWISE, *GO TO Q13*: SAY: Do you think you and your child are able to understand and speak English well enough to participate in this program in English?
    a. *Yes* → *Go to Q13*
    b. *No, cannot speak English well enough to participate* → SAY: I'm sorry to bother you. Thank you for your interest in participating in this study. Unfortunately, at this time we can only do this program in English. I would be happy to talk to you about other resources where your child can get help. Thank you for your time. END CALL.

    SAY: We would like to call you the day before the interview to remind you. Is it OK to call you at this number?
    a. *Yes* → GO TO Q13
    b. *No* → RECORD DIFFERENT NUMBER → GO TO Q12

12. Record new number

    New number: (_____) _____-_____ → Go to Q16

    Say: Thank you for taking the time to speak with me today. End call.

13. Say: Okay, thank you for taking the time to speak with me today. End call. Fill out information below:

    Refusal information

    Who did you speak to? _____

    *Reason for refusal?*_____

4. *Pretest the field instrument*
   Pretesting means identifying a relatively small sample of people who are willing to go through each question with you and tell you what it means to them. (This method is called "cognitive pretesting".) Do the participants agree with your interpretation of each question and response? Usually pretests are done using early versions of the study, and so glitches should be expected. You will need to find a secluded place to conduct the pretest, which is almost always an interview. A trained interviewer is needed. Strict rules are needed for recording participants' answers. Experienced personnel are needed to interpret pretest results and translate them into improvements.
5. *Pilot-test the instrument*
   Pilot testing means having access to a group of potential respondents that is willing to try out an instrument that may be difficult to understand or complete. Expertise is needed in analyzing the data from the pilot test, and experience in interpreting respondents' responses is essential. Additional knowledge is needed in how to feasibly incorporate the findings of the pilot test into a more final version of the instrument.
6. *Administer the instrument*
   Face-to-face and telephone or computer-assisted interviews require skilled and trained personnel. Interviewers must be able to elicit the information called for by the interview questionnaire and record or code the answers in the appropriate way. Interviewers must be able to talk to people in a courteous manner and listen carefully. Also, they must talk and listen efficiently. If the interview is to last no longer than 10 minutes, the interviewer must adhere to that schedule. Interviews become increasingly costly and even unreliable when they exceed their allotted time.
   Among the types of expertise required to put together a mail questionnaire is the ability to prepare a mailing that is user friendly (e.g., includes a self-addressed envelope) and the skill to monitor returns and conduct follow-ups with those not responding. Email surveys also require similar skills. The instrument used to collect data must be user-friendly, and you need the skills to keep track of responses and then follow-up non-respondents.

If you plan to conduct online studies, you should consider becoming familiar with commercial software packages that guide survey preparation and analysis. Training in their use may be necessary for projects that do not have a specialist. If the study is being done at a local site (hospital, clinic), then privacy concerns associated with the Web may be especially daunting.

Expertise is needed in defining the skills and abilities needed to administer the study's field measures and in selecting people who are likely to succeed in getting reliable and valid data. Training is the key a. For example, a poorly trained telephone interviewer is likely to get fewer responses than a well-trained interviewer. Because of the importance of training, many large studies use educational experts to assist them in designing instructional materials and programs for training.

In large and long-term studies, quality must be monitored regularly. Are interviewers continuing to follow instructions? Who is forgetting to return completed interviews at the conclusion of each 2-day session? If deficiencies in the process are noted, then retraining may be necessary.

7. *Manage the data*
   Managing data means programming, coding, and data entry. It also means setting up a database. Programming requires relatively high-level computer skills. Coding can be very complicated, too, especially if response categories are not precoded. Training and computer skills are needed to ensure that data enterers are expert in their tasks. Finally, data cleaning can be a highly skilled task involving decisions regarding what to do about missing data, for example.
8. *Analyze the data*
   Appropriate and justifiable data analysis is dependent on statistical and computer skills. Some studies are very small and require only the computation of frequencies (number and percentages) or averages. Most, however, require comparisons among groups or predictions and explanations of findings. Furthermore, measures of attitudes, values, beliefs, and social and psychological functioning also require knowledge of the statistical methods for ascertaining reliability and validity.
9. *Report the results*
   Writing the report requires communication skills, including the ability to write and present results in tables and figures. Oral presentations require ability to speak in public and to prepare presentations. It helps to have outside reviewers critique the report; time must be spent on the critique and any subsequent revisions. Expenses for reports can mount if many are to be printed and disseminated. ♦

Use the following checklist as a guide in calculating costs and preparing field study budgets.

Costs of Field Work: A Checklist

- Learn about direct costs. These are all the expenses you will incur *because* of the fieldwork. These include all salaries and benefits, supplies, travel, equipment, and so on.

- Decide on the number of days (or hours) that constitute a working year. Commonly used numbers in the U.S. are 230 days (1840 hours) and 260 days (2080 hours). You use these numbers to show the proportion of time or "level of effort" given by each staff member. Obviously these numbers will vary from country to country.
  Example: A person who spends 20% time on the study (assuming 260 days per year) is spending 0.20 × 260, or 52 days or 416 hours.
- Formulate fieldwork tasks or activities in terms of months-to-complete each.
  Example: Prepare instrument during Months 5 and 6.
- Estimate the number of days (or hours) you need each person to complete each task.
  Example: Jones, 10 days; Smith, 8 days. If required, convert the days into hours and compute an hourly rate (e.g., Jones: 10 days, or 80 hours).
- Learn each person's daily (and hourly) rate.
  Example: Jones, US $ 320 per day, or US $ 40 per hour; Smith, US $ 200 per day, or US $ 25 per hour.
- Learn the costs of "benefits" (e.g., vacation, pension, and health) – usually a percentage of salarie.
  Example: Benefits are 25% of Jones's salary. For example, the cost of benefits for 10 days of Jones's time is 10 × 320 per day ×0.25, or US $ 800.
- Learn the costs of other expenses that are incurred specifically for *this* study.
  Example: One 2-hour focus group with 10 participants costs US $ 650. Each participant gets a US $ 25 honorarium for a total of US $ 250; refreshments cost US $ 50; a focus group expert facilitator costs US $ 300; the materials costs US $ 50 for reproduction, notebooks, nametags, and so on.
- Learn the indirect costs, or the costs that are incurred to keep the study team going. Every individual and institution has indirect costs. Indirect costs are sometimes a prescribed percentage of the total cost of the field work (e.g., 10%).
  Example: All routine costs of doing "business," such as workers' compensation and other insurance; attorney's and license fees; lights, rent, and supplies, such as paper and computer disks.
- If the fieldwork lasts more than 1 year, build in cost-of-living increases.
- Be prepared to justify all costs in writing.
  Example: The purchases include US $ 200 for 2000 labels (2 per student interviewed) at US $ 0.10 per label and US $ 486 for one copy of MIRACLE software for the data management program.

10.9 Conclusions

Fieldwork in epidemiological studies involves collecting information to describe, compare, or explain knowledge, attitudes, and behavior about the health and health care of populations. To assure reliable information, field work depends upon asking straightforward questions. Straightforward questions are purposeful, concrete

and expressed as complete questions. Responses may be considered as nominal or categorical, ordinal and numerical. Open questions allow the respondent to give answers in his or her own words. Coding open responses may be difficult. Closed questions provide the respondent with choices. They are easier to interpret and analyze than open questions but may not provide in-depth information. An instrument's length is dependent upon the resources available to develop and validate a questionnaire. Keep in mind that very long instruments may tire some respondents, thereby reducing the reliability and validity of the results. Questions should be ordered logically and each such be related to the expressed purposes of the study. Relatively simple questions should go first, hardest second. Demographic information is often called for in last place.

Make certain that respondents understand the purposes of the study and each question you plan to ask. If questionnaires are to be completed by mail, include self-addressed envelopes. Try to keep questionnaires as short as possible. For online surveys, avoid the need for the respondent to follow many steps: keep the questionnaire short and easy to use. For all self-administered questionnaires, make sure they are pre tested and pilot tested; when possible look at preliminary data to check that all questions are being answered. Interviewing only succeeds with trained interviewers and a method for monitoring the quality of the process. Consider incentives to compensate respondents for their time. Pilot testing is essential to ensure the collection of reliable data. Reliability refers to the consistent with which questions are answered, while validity refers to the accuracy of the answers. Common types of reliability to consider include test-retest and internal consistency. Common types of validity are content, concurrent, predictive, and construct. To improve reliability and validity, check to see that the language and cultural assumptions of the field study are consistent with those associated with the population being studied. Consider using advance letters and incentives to encourage participation and improve response rates. Make certain all measures and the study's logistics are pre tested and pilot tested. Regardless of the methods used to collect data in the field, be ever mindful of the need to ensure confidentiality of responses. Field work tends to be costly because of its dependence upon human capital including trained field workers and data managers.

References

Anastasi A (1982) Psychological testing. Macmillan, New York

Bernard HR (2000) Social research methods: qualitative and quantitative approaches. Sage, Thousand Oaks

Bourque LB, Fielder EP (2003a) How to conduct self-administered and mail surveys. Sage, Thousand Oaks, London, New Delhi

Bourque LB, Fielder EP (2003b) How to conduct telephone surveys. Sage, Thousand Oaks, London, New Delhi

Brett A, Grodin M (1991) Ethical aspects of human experimentation in health services research. JAMA 265:1854–1857
Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20:37–48
Dawson B, Trapp RG (2001) Basic & clinical biostatistics. Lange Medical Books/McGraw-Hill, New York
Fink A (2002) How to ask survey questions. Sage, Thousand Oaks, London, New Delhi
Fink A, Kosecoff JB (1998) How to conduct surveys: a step-by-step guide. Sage, Thousand Oaks, London
Frey JH (1989) Survey research by telephone. Sage, Newbury Park
McDowell I, Newell C (1996) Measuring health: a guide to rating scales and questionnaires. Oxford University Press, New York
Stewart AL, Ware JE (1992) Measuring functioning and well-being: the medical outcomes study approach. Duke University Press, Durham
Sudman S, Bradburn NM (1982) Asking questions. Jossey-Bass, San Francisco