

## Field Methods in Epidemiology

Patricia Hartge and Jack Cahill

---

### Data Collection Instruments

Abstracts • Questionnaires • Physical Examinations • Biospecimen Collection • Environmental Samples

### Field Operations

Experimental Studies • Cohort Studies • Case-Control Studies

### Data Preparation

### Changing Field Methods

---

To succeed in its goals, every epidemiologic study requires sound design, sensible analysis, and careful execution. Execution of the study, or field work, ought to be guided by a simple principle: Create the strongest bridge between the design and the analysis. To accomplish this, the epidemiologist tries to select or devise valid and reliable data collection instruments, field operations, and office procedures, attending especially to elements of the design essential to validity.

Field work demands so much time in all but the smallest studies that the epidemiologist responsible for the study seldom does all of it, relying instead on study staff. In medium-sized studies, the daily operations of the study staff are directed by a field supervisor familiar with field methods but not necessarily trained in epidemiology. In large, multi-center studies with a field staff of dozens, each center may have its own study manager as well. An experienced and capable study supervisor provides enormous benefit to the study but does not relieve the epidemiologist of responsibility for field work.

Epidemiologists can ensure the quality of the field work in many ways. They ought to oversee the design of the data collection instruments after consulting relevant examples and literature inside and outside of epidemiology. Whenever possible, epidemiologists ought to test the data collection instruments themselves during the design phase, even if they are not professional interviewers or abstractors. They should document the field operations and office procedures in manuals used for training and reference. They should incorporate quality-control methods in each phase of the study. In addition, they ought to conduct substudies of reliability and validity to improve the quality of the field work and aid in the analysis and interpretation of findings.

### DATA COLLECTION INSTRUMENTS

Epidemiologists seldom collect data without written, standard instruments designed to increase the comparability between the study groups. Epidemiologic studies use a vari-

ety of instruments: record abstracts, in-person interview questionnaires, self-administered questionnaires, telephone interview questionnaires, physical examinations, biospecimen collections, and environmental samples. Responses can be recorded either on paper or electronically.

Most of the instruments need to be pretested on a small sample and revised and tested again in a pilot study. The intended users—study subjects, interviewers, abstractors, examiners, and so forth—ought to evaluate the instruments. The longer, more complex forms need further documentation with item-by-item specifications of the intent of the question and any peculiarities or exceptions. These specifications, incorporated into one of the study manuals, aid in staff training and serve as a reference after data collection ends.

### Abstracts

Abstracts are used to distill information needed for the study from written records kept for other purposes, such as medical charts or employment records. The larger and more variable the original record, the harder the abstract is to design. The designer tries to make the layout clear, the wording consistent, and the path through the abstract evident from the shading, indentation, arrows, etc., especially for items that depend on an answer to another item. As much as possible, the form ought to use closed-ended questions, those with a prerecorded set of responses from which to select. Open-ended questions may be used to capture uncommon or more detailed responses. The form may require the abstractor to distinguish between negative findings and absent findings for important items. For instance, some items can be grouped into one list, with an instruction to indicate all that apply; but if it is important to know that the abstractor looked and did not find mention of a particular clinical finding, then the question ought to be formatted so that an answer must always be indicated (yes/no/not found).

Medical record abstracts can be hard to design because medical charts are complex and idiosyncratic. Medical record abstractors do receive training in the medical terminology pertinent to the study, but they cannot be expected to glean everything from the charts that a physician specialist would. The abstract form ought to be designed to reduce the need for interpretation, even at the expense of collecting items that may be redundant (Fig. 11-1).

Even if expert medical personnel will abstract the records, practices vary among hospitals and clinics, so the investigator may need to review typical and atypical charts in the study sites before designing the abstract form. The designer also may need to interview some recording physicians or other health personnel to understand the recording practices. When the analysis will exclude events before or after particular dates, the form may be designed to capture only the relevant period or the entire history, with exclusions made in the analysis.

Devising good abstract forms may be hard, but testing them usually is easy. The investigator abstracts several records, and then one or more of the abstractors independently abstract the same records. This exercise can reveal gross errors or oversight in form design. A similar process also measures interabstractor reproducibility during form development.

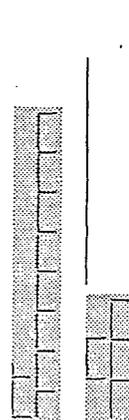
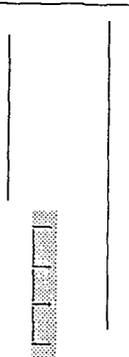
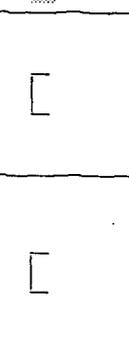
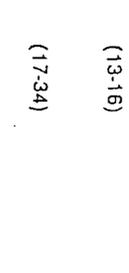
### Questionnaires

Questionnaires remain the mainstay of epidemiologic data collection. A recent survey of case-control studies found that 67% used a questionnaire as the sole source of exposure data and an additional 16% used a questionnaire and other sources (Correa et al., 1994).

NEOPLASMS

REC.06

11. MALIGNANT AND BENIGN NEOPLASMS OTHER THAN RB: 0 = No 1 = Yes 7 = Unclear/Discrepant 9 = Not mentioned.....  (12)  
 (If Yes, record details in table below; otherwise, skip to Item 12)

NEOPLASM TABLE		DATE OF DIAGNOSIS	PLACE DIAGNOSED	CONFIRMATION SOURCE	PHOTOCOPY OF CONFIRMATION
(Exclude RB metastases or recurrences)		MO	(Hospital/Physician Address)		0 = Not in Record 1 = Yes
a.				<input type="checkbox"/>	<input type="checkbox"/> (13-16)
				<input type="checkbox"/>	<input type="checkbox"/> (17-34)

CONFIRMATION SOURCE: 1 = Surgical Pathology Report 3 = Radiology Report 5 = Clinical Only 8 = Other (SPECIFY)   
 2 = Autopsy Summary Report 4 = Death Certificate Only 6 = Correspondence Only 9 = Not Mentioned

FIG. 11-1. Sample item from medical abstract used in a retinoblastoma (RB) study.

Questionnaire design, constrained by some of the same requirements as abstract design, also grapples with the issues of cooperation, fatigue, meaning, memory, and honesty. Errors in interview data collected by questionnaire can be introduced by the respondent or by the interviewer. Questionnaire design aims at reducing both. Design of epidemiologic questionnaires benefits from a vast body of survey research experience and literature (Biemer et al., 1991; Groves, 1989; Sudman and Bradburn, 1991; Tanur, 1991).

### *Questionnaire Methods*

The first design choices concern the administration of the questionnaire: Who ought to record the answers to the questions—the subject or an interviewer, and where should the subject be—at home, in the clinic or hospital, or elsewhere? If the questions are short and simple and the respondents are willing and able to complete a questionnaire, a self-administered questionnaire is feasible (see Fig. 11-2).

The self-administered questionnaire may be mailed to the subject at home or given in person. Self-administered instruments generally cost less than interviewer-administered instruments, but they also are the least easily monitored and the most susceptible to misunderstood questions and skipped answers. If only a modest fraction of the items are left blank, study staff can ask the respondent for clarification in person or by telephone, keeping labor costs lower than they would be with interviewer administration.

Self-administered questionnaires offer advantages other than cost. They may yield more accurate data on sensitive or embarrassing topics because they are more anonymous. Because printed visual aids can be incorporated, some topics are measured more easily in self-administered questionnaires than in telephone interviews. Self-administered questionnaires generally require a printed paper format. Optically scanned forms may be considered if the answers are expressed as multiple choices. For smaller studies, the cost of developing the scanned forms may exceed the savings in coding and keying.

If the questions are complex or nested (with different answers leading to different series of questions) or if the answers are likely to need probing for clarification, then an interviewer ought to administer the questionnaire either in person or by telephone. The age, education, health, and hearing of the subjects, the budget, and the number and types of questions contemplated all influence the choice between telephone and in-person interviews.

Personal interviews generally entail more labor cost than do telephone interviews, because of unavoidable travel and appointment time between one interview and the next. The greater intimacy of in-person interviews may increase the subject's willingness to participate in the study. Personal interviewing typically increases the length of individual responses and of the total interview, but whether this increases accuracy is unknown (Groves, 1989).

Personal interviews can incorporate pictures, three-dimensional models, and other memory aids. If the questionnaire requires the respondent to recall distant events, such memory aids may be critical. For instance, photographs of medications help the respondent recognize individual formulations. Food models describe portion sizes in dietary interviews. Maps are used in questions pertaining to residence or travel. Diaries, timelines, or calendar grids can improve completeness and dating of lifetime residential, occupational, or reproductive histories. If there are no published evaluations of the contemplated memory aids in a similar setting, a small pilot test of such aids may help determine whether they are effective and feasible.

tract design, honesty. Respondent or epidemiologic and literature (1).

Who ought to be should the are short and re, a self-ad-

e or given in administered otible to mis-items are left phone, keep-

ny may yield more any- asured more self-adminis- anned forms smaller stud- coding and

ifferent series then an inter-. The age, ed- types of ques- interviews. interviews, be- and the next. willingness to of individual is unknown

ls, and other t events, such p the respon- in dietary in- ies, timelines, ntial, occupa- contemplated elp determine

10. Have your menstrual periods stopped, that is, have you been through menopause?

1  Yes, periods stopped

2  Had menopause but now have periods due to hormone therapy

3  No, still menstruating

4  Not sure

Explain \_\_\_\_\_

GO TO QUESTION 11

GO TO QUESTION 11

10a. At what age did your periods stop? \_\_\_\_\_ AGE

10b. What is the reason your periods stopped?

1  Surgery

2  Natural menopause

3  Other reasons - Specify \_\_\_\_\_

10c. Has your uterus (womb) been removed? [If yes, indicate year of surgery and reason for removal.]

	<u>Year of Surgery</u>	<u>Reason for Removal</u>
1 <input type="checkbox"/> Yes.....	.....19__	1 <input type="checkbox"/> Fibroid tumor
		2 <input type="checkbox"/> Abnormal bleeding
2 <input type="checkbox"/> No		3 <input type="checkbox"/> Ovarian cysts
		4 <input type="checkbox"/> Endometriosis
		5 <input type="checkbox"/> Prolapsed (dropped) uterus
		6 <input type="checkbox"/> Other - Specify _____
		8 <input type="checkbox"/> Don't know

FIG. 11-2. Example of a self-administered questionnaire.

Telephone interviews typically elicit slightly shorter answers than do personal interviews (Groves 1989). Respondents also favor the first answer when a list of possible answers is read to them on the telephone, but other differences between the telephone and personal interviews are modest (Groves, 1989). Supervisors can monitor interviewing more easily in telephone interviews than personal interviews because they can listen to the conversation. Some sensitive topics can be difficult to query by telephone, because the respondent may become suspicious of the interviewer or the legitimacy of the study. On the other hand, once the respondent trusts the study and the interviewer, questions about socially undesirable behaviors may be answered more readily because of the greater social distance in a telephone call (Sudman and Bradburn, 1991).

The overarching consideration in deciding on the methods of questionnaire administration should be comparability between study groups. Thus, case-control studies comparing hospitalized cases and controls lend themselves to personal interviews in the hospital. On the other hand, if cases derive from selected hospitals and controls from neighbors or the general population (to approximate the referral network), then interviews would be most comparable if they are conducted with all subjects at home. If the

various aims and constraints of the study suggest more than one questionnaire mode, then the investigator may conduct a pilot study comparing them.

Most questionnaires in epidemiology are printed, but increasing numbers are automated (computer-assisted personal or telephone interviewing, CAPI or CATI). Computerization eliminates the time between data collection and access to the data for analysis. For instruments with complex logical branching, computerization reduces interviewer error. Costs of developing and deploying such instruments continue to fall but still prohibit their use in smaller studies.

### Questions

After choosing the type of questionnaire, the investigator focuses on framing the questions. This topic has gotten an appropriately large amount of attention in survey research literature (e.g., Sudman and Bradburn, 1991; Clark and Schober, 1991). Wording matters in all surveys because people construe important and simple words like *anyone*, *most*, *average*, *never*, or *fairly* in varied ways (Groves, 1989). Medical terms with precise meanings need to be explained simply, so "people related to you by blood" might be preferable to "family" and "Has a doctor ever told you that you had . . . ?" might be preferable to "Have you had . . . ?" On the other hand, vernacular expressions are not always preferred (e.g., "medications" ought to be used instead of "drugs" to avoid confusion with illicit drugs).

Question length also affects response. Short questions usually are clearer, but a longer question may improve respondent cognition. For instance, it is hard to absorb all the details of the question, "How many hours per week do you typically use this product in the summertime on weekdays?" The question may be more intelligible if it can be separated into specific questions to reduce the density of concepts (Tanur, 1992) or preceded with a description of some of the concepts (e.g., "I'd like you to think about the summertime. I will be asking you about weekdays separate from weekends.>").

How many questions does the investigator need to ask on a topic? One or two questions may not gather enough information to analyze thoroughly, particularly if an unexpected association emerges. On the other hand, a long series of questions on a single topic bores the respondent and results in missing or inaccurate answers after the first few questions. Assessments of usual diet, lifetime occupational history, and lifetime residential history each take 10–20 minutes and are particularly prone to this problem.

The order of questions can also affect response. Many questionnaires begin with questions that are not threatening and not taxing to the memory and, if possible, interesting to the respondent (Sudman and Bradburn, 1991). Questions about sensitive topics usually follow related but less sensitive topics. Some instruments precede them with a prologue that acknowledges the personal nature of the question, reiterates the subject's right not to answer, and states the importance of the question to the survey. The format of answers matters, too. For open-ended numerical responses (e.g., how many years), respondents often show digit preference (for 0 and 5). For closed-ended items with an ordered list (e.g., <5, 5–10, 11–19, ≥20), some respondents may favor a part of the range, often the central part.

How much information can the questionnaire elicit? Personal interviews commonly last 50–90 minutes; telephone interviews typically last 30–60 minutes; self-administered questionnaires typically take 10–20 minutes to complete. In our experience, interviewers and respondents report these times to be acceptable, and common sense suggests much longer times are not feasible. The length of the interview does little to change the willingness of the study subject to agree to the interview. Usually, the respondent continues with the in-

terview to th  
(Groves, 19  
possible, ev  
lengthens. S  
the effects c

Memory  
posures unc  
they happer  
ond, many  
improve re  
representat  
does not al  
in an invest  
tween case  
devices are  
stimulate r

Whether  
spondents  
ries in rev  
on the exp  
ting is to t  
of responc  
porting (L

Special  
If the que  
beginning  
that occur  
"Since we

Some r  
groups) v  
many oth  
keep the i

Standar  
stance, a  
jects who  
spondent  
comparal  
collection  
for nonre  
assessed  
are expe  
stantially  
complete

Epid  
count no

then interview to the end. On the other hand, longer interviews may increase the risk of breakoffs (Groves, 1989). Most epidemiologists find it hard to resist including as many questions as possible, even if some degradation in the quality of the answers occurs as the interview lengthens. Some researchers place least critical data near the end in case of breakoff, but the effects of this practice are not known.

Memory presents the hardest problem in epidemiologic questionnaires. First, many exposures under study did not seem interesting or important to the respondent at the time they happened, let alone decades later (Croyle and Loftus, 1991; Loftus et al., 1991). Second, many of the cognitive interviewing techniques (Fisher and Quigley, 1991) that can improve recollection dramatically (e.g., context reinstatement, focused retrieval, multiple representation) do not suit epidemiology. Standardization of instruments and techniques does not allow interviewers to pursue a topic as they would in an ordinary conversation or in an investigative interview, but standardization does keep data collection comparable between cases and controls or exposed and unexposed. Nonetheless, some recall-improving devices are applicable to epidemiology, such as using benchmarks in the subject's life to stimulate recall of a time period.

Whether to use chronologic or reverse chronologic order depends on the exposure; respondents may recall reproductive histories in chronologic order, but occupational histories in reverse. Whether to ask "How old were you..." or "In what year..." also depends on the exposure. For many exposures, previous studies may suggest what level of forgetting is to be expected. Furthermore, when historical records can be obtained on a sample of respondents, they may provide some measure of the accuracy of recollection and reporting (Linet et al., 1989).

Special considerations arise with longitudinal studies that require repeated interviews. If the questionnaire arrives on a strict annual schedule tied to the subject's birthday or the beginning of each calendar year, it may succeed in getting subjects to report only events that occurred since the last contact. Otherwise, it is best to avoid questions that refer to "Since we asked you last..." because of the risk of events being forgotten or repeated.

Some multipurpose studies (e.g., a study with one control group and multiple case groups) use a modular questionnaire. A few core questions are asked of everyone, and many other questions are asked only of a subset of the respondents. This approach can keep the interview to a reasonable length.

Standardization must be compromised at times to avoid losing data completely. For instance, a few telephone interviews may be conducted in a personal interview study for subjects who would not participate otherwise. If a different mode, source (e.g., surrogate respondent), or setting is used, this fact must be recorded. Despite efforts to maintain comparability with standard instruments and training, differences in the quality of the data collection may arise. To detect such differences, one may compare the rates and reasons for nonresponse, the length of the interview, the quality of understanding and response as assessed by the interviewer (poor, adequate, excellent), and the responses to questions that are expected not to differ between the groups (exposures or outcomes). For example, substantially longer interviews in one group raises the possibility of differential accuracy or completeness of information.

### Physical Examinations

Epidemiologic studies can include physical examinations to measure blood pressure, count nevi, assess body fat distribution, and so on. The order and content of these exams

is much more explicitly prescribed than in a strictly clinical setting, and the exams also require more exact recording of findings. Many subjects with no abnormal findings whatsoever (including most controls in case-control studies) must be fully examined and described to ensure comparability in the epidemiologic study. If any of the findings is abnormal, the subject must be referred for clinical evaluation.

Exam forms often have complex logical branching, and the examiner has to be able to follow the flow and complete the form during the exam. Exam forms usually need extensive pretesting to make sure they are easy to use during the exam. Variation among clinical observers remains a concern even with standardized instruments, so training and quality control are especially important. For example, training in blood pressure measurement often uses a videotape test as an objective standard. Studies of interobserver variation may be necessary when the physical exams are subtle or complex.

Physical exams in epidemiologic studies may be conducted by medical or paramedical personnel, with oversight by a physician or other expert reviewer. If the expert reviewer does not see each patient, then the examiner ought to record whether expert review occurred and distinguish the expert's evaluation from the original one.

### Biospecimen Collection

Epidemiologic studies increasingly involve laboratory components requiring collection of urine, blood, or tissue from subjects. The biggest problems that arise often reflect unrecognized laboratory error. Indeed, epidemiologists usually need to investigate the reproducibility of the assay in a pilot study. Field problems also may arise in applying tests or assays to large-scale epidemiologic studies that previously have been conducted only on a small scale. For instance, laboratories that have developed an assay and can achieve high reproducibility with dozens of specimens may be unable to process the thousands of specimens required in a field study. Another typical problem arises when the investigator wishes to collect specimens for storage without a definite assay in mind, so the particular collection and storage requirements are not certain.

Nurses or phlebotomists usually collect the specimens. The protocol for collecting, processing, labeling, storing, and shipping the specimens ought to be documented in the manual of operations. The biospecimen collection forms record the subject, the specimen number, whether baseline or subsequent, results of the collection (e.g., number of tubes drawn, medical complications), processing and storage (e.g., which freezer). Bar-coded labels reduce handwriting or keying errors, and clear, detailed shipping and storage lists minimize losses in transit and aid in tracking those that occur.

### Environmental Samples

Epidemiologic studies may include taking measurements of water, soil, air, ionizing radiation, electromagnetic fields, and other environmental variables. Combining these assessments with the individual history of occupation, residence, etc., becomes complicated if the levels of exposure at a particular place have changed over time or if the subject has moved or changed jobs, so that multiple old sites need to be monitored. Usually, the individual data are linked to the environmental data at analysis, and environmental data forms are designed to capture the data from one site, not linked to a par-

ticular individual  
der which the

Each of the  
Experimental  
dures, interven  
and acquisition  
death certifica  
view), exposure  
ministration,  
Case-control  
administration

Experimen  
teers (field tr  
munity trials  
ing, involve  
recruitment :  
any trial that  
particular de  
Recruitment  
so that study  
designed to  
identified pe  
gists and by

Without li  
releases, and  
prominent p  
recruitment  
(e.g., a free  
or other min  
mands of the  
sation may l  
dollars) do  
payments ca  
needed in st

Once recr  
the trial by  
ization can  
groups, the  
of the main  
ticipants un  
Many interv

g, and the exams also  
o abnormal findings  
st be fully examined  
y. If any of the find-  
ion.

miner has to be able  
1 forms usually need  
the exam. Variation  
ized instruments, so  
le, training in blood  
standard. Studies of  
is are subtle or com-

medical or paramed-  
wer. If the expert re-  
rd whether expert re-  
iginal one.

nts requiring collec-  
is that arise often re-  
ly need to investigate  
also may arise in ap-  
reviously have been  
ave developed an as-  
ns may be unable to  
ther typical problem  
ge without a definite  
its are not certain.

otocol for collecting,  
be documented in the  
he subject, the speci-  
tion (e.g., number of  
, which freezer). Bar-  
ed shipping and stor-  
ccur.

ter, soil, air, ionizing  
les. Combining these  
, etc., becomes com-  
ed over time or if the  
eed to be monitored.  
at analysis, and envi-  
e, not linked to a par-

ticular individual. On the form, the sampling technician also records the conditions under which the monitoring occurred, as well as the levels detected.

## FIELD OPERATIONS

Each of the main types of epidemiologic studies requires characteristic field tasks. Experimental studies typically involve volunteer recruitment, randomization procedures, intervention activities, questionnaire administration, matching to mortality files, and acquisition of data on health outcomes (medical records, pathology reports, and death certificates). Cohort studies involve cohort assembly (usually from record review), exposure classification (usually record abstraction), tracing, questionnaire administration, matching to mortality files, and acquisition of data on health outcomes. Case-control studies involve selection of cases, selection of controls, and questionnaire administration (to assess exposures).

### Experimental Studies

Experimental studies assign exposures randomly to patients (clinical trials), volunteers (field trials), or whole communities (community trials). Clinical trials and community trials will not be described here. Field trials, whether of prevention or screening, involve the same field tasks as cohort studies do but also the tasks specific to recruitment and randomization. Participants need to be convinced of the benefits of any trial that makes considerable demands on the subjects. If the trial is targeted to a particular demographic group, this target may suggest possible recruitment sources. Recruitment is easiest if lists of names and phone numbers or addresses are available so that study staff can approach prospective recruits individually. For example, in a trial designed to investigate diet and the recurrence of adenomatous polyps, investigators identified potential participants by getting referrals from participating gastroenterologists and by reviewing medical records of participating endoscopy services.

Without lists of potential recruits, investigators must recruit by public notice, news releases, and advertising. Investigators often solicit sponsorship or endorsement by prominent people and community or medical organizations to increase interest. Most recruitment succeeds by persuading the subjects of the value of the trial to themselves (e.g., a free medical exam) and to society. Some trials offer reimbursement for parking or other minor costs of participating in the trial. If the time commitment or physical demands of the study are great (e.g., with multiple blood collections), financial compensation may be necessary. In our experience, cash incentives of less than \$25 (current dollars) do not increase participation very much. On the other hand, somewhat larger payments can affect participation. Payments over \$100 are uncommon but may be needed in studies that make many demands on the participants.

Once recruited to participate, subjects are randomly assigned to one of the arms of the trial by use of a random number table (paper or computerized). Because randomization cannot guarantee the similarity of the participants among the experimental groups, the investigator ought to use a baseline questionnaire to measure the predictors of the main outcomes. The investigator also collects information for locating the participants until the end of the follow-up period (e.g., names and addresses of friends). Many intervention trials require annual or other regular follow-up by mailed question-

naire, with telephone calls to remaining nonrespondents. Even if no new data are needed on outcomes or exposures, addresses need to be kept current.

### Cohort Studies

The study protocol defines the cohorts of subjects to be studied. The operations manual provides detailed definitions of eligibility, instructions for finding the eligible subjects, and instructions for getting access to the subjects or their data. If a cohort study requires collecting details of the exposure (timing, intensity, other exposures), the same data sources may be needed to assemble the cohort and characterize the exposures of individual members. Studies of occupational and medical cohorts typify the field methods used in cohort studies in general.

#### *Cohort Assembly and Exposure Classification*

In an occupational cohort study, the investigator assembles the cohort from the records of a company, a union, or a professional or trade association. Many preliminary studies use union or association records alone. These often permit assembly of a complete cohort but lack the detail on tasks and locations of work essential in defining each individual's jobs and exposures. If both union and company records are available, both sources may be used to increase completeness.

At the outset, the study team (typically an epidemiologist, an industrial hygienist, a study manager, and one or more abstractors) visits some of the plants and the headquarters or offices in which worker records are kept. The investigators inquire about every possible source of records, including occupation records, W-2 forms, payroll ledgers, union rolls, medical records, life and health insurance systems, and Social Security Administration and Internal Revenue Service records, both computerized and paper. Although the separate record systems will be incomplete (Marsh, 1982), together they may provide a nearly complete enumeration of the cohort.

To uncover as many record sources as possible, the investigator interviews many potential informants at different levels of authority, from clerical to managerial. It is necessary to ask about record systems no longer maintained and about groups of records stored separately, such as those from pensioners, workers terminated before pension eligibility, workers involved in litigation, or workers under medical care. Once the records are found, research staff abstract or photocopy them. Photocopying costs more but allows the investigators to see the records at the office, check on the quality of abstracting, and glean additional data. A modest cohort of 5000 workers typically has held 50,000 job lines (the combination of job title and department). Research staff may bring their own copiers to avoid disruptions of ordinary business.

At the study offices, the study manager oversees the compilation of the abstracts or photocopies of the various records into the preliminary cohort. Almost always, this preliminary cohort is not complete because some records were lost or overlooked. Such losses often are detected by scanning the frequency distributions in dates of hire or termination, first letter of last name, occupational status at termination, branch or plant location, job type, and duration of employment. Staff recollections of numbers of employees and duration of employment are notoriously unreliable, generally from remembering long-term more than short-term workers. After searching for any possible gaps, the in-

vestigators re-  
assembly req

The research  
partments) ar  
mon tasks an  
view the inc  
personnel to  
jobs, based o  
ies, the third  
vironmental :

Studies of  
seen in occup  
horts, group  
medical treat  
files, pharm  
ity logs kept  
surgery, or o  
medical rec

Apart from  
posure often  
able. Investig  
tice to invest  
to determine  
diagnosis an  
sure, such as

Some stud  
status, inclu  
cates can b  
Death Inde  
Cost and co  
Kraut et al.

If study s  
follow-up),  
name, addr  
spouse's na  
last address  
investigato  
vital status

Subjects  
and the U.  
reau files (c  
credit) and  
at the date  
not indicat

investigators revise the cohort. Straightforward in principle, complete occupational cohort assembly requires considerable effort.

The research staff first captures the work history (progression through job titles and departments) and then classifies all job title/department combinations into jobs with common tasks and locations. The abstractors, supervised by the study manager, typically review the individual work histories from records and conversations with company personnel to resolve discrepancies. The industrial hygienist collapses the job lines into jobs, based on familiarity with the work environment (Stewart et al., 1992). In some studies, the third task is to impute exposure levels to jobs, based on current and historical environmental samples and details of job activities provided by workers and supervisors.

Studies of cohorts selected outside of work settings present many of the same issues seen in occupational-cohort studies. Some particular issues arise in studying medical cohorts, groups of people whose exposure of interest is a disease, medical condition, or medical treatment. Study staff select cohort members from hospital discharge diagnosis files, pharmacy records, medical insurance data, birth certificates, or from routine activity logs kept by medical practices, clinics, and hospital departments, such as pathology, surgery, or obstetrics. Cohort assembly encounters problems because some of the needed medical records have been destroyed, lost, or stored in inconvenient locations.

Apart from logistical problems in obtaining the medical data, the classification of exposure often presents the greatest problem because medical records are complex and variable. Investigators generally make several preliminary visits to the hospital, clinic, or practice to investigate the sources and quality of data. Multiple record sources may be needed to determine whether a subject is eligible (e.g., surgical pathology logs to determine the diagnosis and hospital patient files to obtain the demographic data). An uncommon exposure, such as diethylstilbestrol, necessitates the review of many records (Nash et al., 1983).

#### *Tracing and Follow-up*

Some studies do not require contact with the subject but only a determination of vital status, including date and place of death for those who have died, so that death certificates can be gotten. Major United States vital status data sources include the National Death Index, Social Security Administration files, and Veterans Administration files. Cost and coverage vary among them (Boyle and DeCoulfe, 1990; Calle and Terrell, 1993; Kraut et al., 1992; Williams et al., 1992).

If study subjects are to be queried by questionnaire on outcomes or exposures (active follow-up), they first must be located. With recent, accurate, detailed information on the name, address, telephone and social security numbers, parent's names for a child, and spouse's name and employer, the investigator may begin by mailing information to the last address or calling the subject by telephone. Without such detailed information, the investigator uses a variety of methods to locate the subject. If many subjects have died, a vital status search may be the first step.

Subjects not known to have died can be traced in various ways. Municipal directories and the U. S. Postal Service correction service can be used to locate a subject. Credit bureau files can show that the respondent was alive (e.g., at the date of an application for credit) and can provide an address. Driver's license files can show the subject was alive at the date of last renewal and give an address. Car registration and telephone listings do not indicate vital status, since family members can maintain these without removing the

decedent's name. Health Care Financing Administration files contain names, ages, and address of people eligible for Medicare.

The more out of date the last known address, the more difficult the tracing. On the other hand, tracing is easier if many cohort members have died, especially after 1979 when the National Death Index began. The algorithm for tracing depends on the composition of the cohort. In occupational cohort studies in the United States, the investigator often knows the social security number, which helps in matching to mortality, credit bureau, and other files. With medical cohorts (Griem et al., 1994; Inskip et al., 1990), identifying data often is sparser and less accurate, so many sources may be needed to find a large proportion of the cohort. Additional problems arise in following cohorts of women (e.g., from changes in surname) (Boice, 1978) or cohorts exposed before or at birth (Nash et al., 1983).

For studies requiring active follow-up, tracing the subject and collecting the outcome data usually are done in tandem. Procedures for mailing the questionnaires or interviewing by telephone or in person are similar to those in case-control studies. One typical procedure is to mail the follow-up questionnaire once, wait a few weeks for its return, mail again to nonrespondents, and then telephone remaining nonrespondents. The second mailing and the telephone call both increase response markedly in most studies. Typical results of such a follow-up are shown in Fig. 11-3.

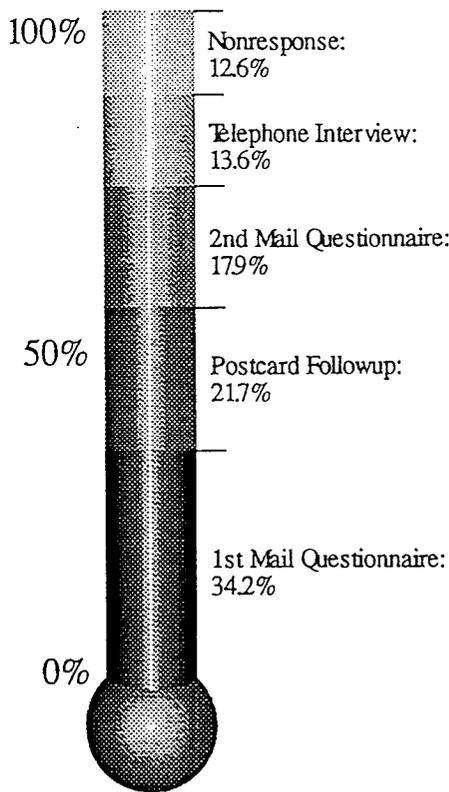


FIG. 11-3. Response rate with each contact in a cohort study. (From Brinton L. *Breast Cancer Detection Demonstration Project, 1987-1989*. National Cancer Institute.)

Longitudinal stu  
some circumstance.  
These subsequent v  
subjects inclined n  
cohort has been fo  
Letters following e  
help keep subjects  
not contacted in su  
multiple opportuni

Often the study  
kin on follow-up c  
of death is known  
United States, ob  
form and takes be  
quire supporting i  
of subjects are usu  
number but know  
Serious illness  
For cancers amon  
able to confirm a  
For other conditio  
tal with permissi  
United States do

A case-control  
the protocol desc  
are developed. Th  
diagnosing it, the  
both etiologic res  
based disease re  
useful sources of

TABLE 7

1986	7
1989	9
1993	9
Cumulative in 1993	(

\*Excludes decea  
From Russell M.  
bany: New York Sta

Longitudinal studies require repeated follow-ups and present special problems. In some circumstances, nonresponders in one wave can be approached in subsequent waves. These subsequent waves may achieve a greater response than the initial follow-up, since subjects inclined not to participate usually refused the first time. On the other hand, if the cohort has been followed several times, motivation to answer more questions may flag. Letters following each round of follow-up describing findings and progress of the study help keep subjects willing to participate. In some studies, once subjects refuse, they are not contacted in subsequent waves. Even with excellent response rates in each wave, the multiple opportunities for nonresponse lower the cumulative response rate (Table 11-1).

### *Confirmation and Details of Morbidity and Mortality*

Often the study requires confirming outcomes reported by respondents or their next of kin on follow-up questionnaires with data from other sources. For instance, once the fact of death is known, the cause of death can be collected from the death certificate. In the United States, obtaining death certificates from states requires a detailed application form and takes between 6 and 12 months, depending on the state. Some states also require supporting institutional review board (IRB) approval letters or study protocol. Lists of subjects are usually grouped into those with known death certificate number, unknown number but known date of death, and unknown date.

Serious illnesses also can be confirmed and detailed by obtaining additional records. For cancers among residents of an area covered by a tumor registry, the registry often is able to confirm a cancer reported by the subject or next of kin (Wasserman et al., 1992). For other conditions requiring hospitalization, records can be requested from the hospital with permission from the patient. Timing is critical because some hospitals in the United States do not honor permission forms more than 60 days old.

### *Case-Control Studies*

A case-control study ought to be identified with an underlying source population, with the protocol describing the population as explicitly as possible before field procedures are developed. The source population reflects the disease under study, the difficulties in diagnosing it, the routine procedures for recording its occurrence, and its frequency. As both etiologic research and computerization of medical data have expanded, population-based disease registries have proliferated. Hospitals and clinics remain convenient and useful sources of cases, despite the challenge of understanding and sampling the under-

TABLE 11-1. *Response rates in repeated follow-ups in a longitudinal study*

1986	78% respond		22% no response	
1989	92% respond <sup>a</sup>	8% no response		No contact
1993	93% respond <sup>a</sup>	7% no response	42% respond <sup>a</sup>	58% no response
Cumulative in 1993	$(.78 \times .92 \times .93) + (.78 \times .08 \times .42) = 69\%$		$.22 + (.78 \times .92 \times .07) + (.78 \times .08 \times .58) = 31\%$	

<sup>a</sup>Excludes deceased, movers, and untraceables prior to field period.

From Russell M. *Stress and Hypertension Study Conducted by the Research Institute on Addiction*. Albany: New York State Office of Alcoholism and Substance Abuse, unpublished data.



1984; Potthoff, 1994) has become one of the most common techniques. Answering machines and fax and data lines have added to the complexity of telephone sampling, and response rates to all types of surveys have been falling for several years (Groves, 1989, page 155). For these reasons, the utility of random digit dialing may diminish. Thus far, variation in response rate among studies has exceeded the temporal decline, in our experience (Table 11-2).

In studies based in hospitals, clinics, or medical practices, investigators gain access to a convenient list of potential cases (e.g., from records in a hospital) not linked to an enumerated underlying population. They must impute the source population and devise a control group to approximate it. Some studies compare hospital-based cases to controls selected from the neighborhoods of telephone exchanges of the cases on the theory that such controls would be referred to the same hospitals if they developed the disease. Another common strategy is to select controls from patients in other clinics within the same hospital. The protocol may specify which diagnoses are to be excluded from the control group because they are related to the exposures of interest. Additional exclusions often include psychiatric diagnoses or other conditions that compromise data collection. One design variation designates a few diagnostic categories to include rather than select from an unspecified mixture of diagnoses.

Procedures for case and control selection in hospital studies should be as precise as possible and remain constant from day to day and patient to patient. Study staff assigned to select a patient from a clinic on a certain day, for example, should be following an algorithm that dictates which particular patient ought to be picked. Such practices not only guard against deliberate or unconscious enrollment of particular patient types but also allow routine quality checks. Data on patients unavailable or unwilling to be studied ought to be recorded in order to characterize nonresponse.

In hospital and population studies, details of subject selection need to be specified in the procedures manual: the source of the diagnostic information (admission, inpatient, discharge, etc.), the algorithm for selection according to date, age, sex, etc. (affected by matching or other stratification schemes in the protocol), any requirements for verification or confirmation of record data, any procedures for permission to interview or abstract data, and the procedures for recording the patient selection.

### Interviewing

Errors introduced by the interviewers can be reduced by good hiring practices, training, and supervision. Interviewers are trained to apply standard interviewing techniques to the particular questionnaire, first with formal training sessions typically lasting 2–3 days and then with days or weeks of practice interviews. For a new interviewer, the in-

TABLE 11-2. *Random digit dialing: response rates in selected U.S. studies of cancers of the female reproductive tract*

Study	Years	Ages	Places	Response rate
Cancer and steroid hormone (CASH)	1980	20–54	8 areas	87
Cervical cancer	1982	18–64	5 areas	84–87
Asian-American breast	1985–1986	20–60	California	91
Rare reproductive	1986	20–64	New York	82
Endometrial cancer	1987–1990	20–64	5 areas	86

interview supervisor checks many of the initial interviews, for instance, by attending the interview or reviewing an audio or video recording of it. After the interviewer has mastered the techniques, a fraction may be audited at random throughout the study.

The interviewer's first job is to persuade the subject to participate. The introduction needs to be framed carefully and pretested to convey the scientific importance of the study without making subjects with particular histories more or less likely to respond. The reputation of the scientific institution sponsoring the study usually convinces the subject of the legitimacy of the study and often its value, but the individual investigator must be available by telephone to vouch for the study. The introduction must accurately portray the time involved. Individual Institutional Review Boards (IRBs) impose additional requirements.

If the subject refuses at first, the interviewer ought to make another effort to persuade the subject or to learn the reason for the refusal. The interviewer may ask the subject to talk to the supervisor or someone else before the subject makes a firm refusal. Interviewers document nonresponses with notes that help the supervisor approach the subject. Especially persuasive interviewers serve as refusal converters and often raise response rates by several points. Whether converted refusers differ from initial responders is not known.

After obtaining cooperation, the interviewer proceeds with the interview or arranges an appointment. Before an in-person interview, the interviewer arranges any equipment, memory aids, or materials. The interviewer tries to arrange the setting and timing to minimize interruptions and distractions to the respondent. Occasionally, interviews must be stopped and resumed later. These and other deviations in the administration of the questionnaire must be noted by the interviewer, as well as an assessment of the quality of the responses.

Questions that the respondent does not understand are repeated slowly verbatim, not rephrased. Further clarification follows the instructions to interviewers given in training and manuals. Answers that are not clear or complete are probed with neutral, standard questions also covered in training and manuals (Table 11-3). To conclude the interview, the interviewer quickly scans the questionnaire for omissions, thanks the respondent, and explains that the supervisor may call the respondent to review the interviewer's work. In some studies, the investigator offers the respondents the opportunity to be notified of the findings.

The study supervisor reviews completed questionnaires, querying the interviewer about any unclear entries. In many studies, the supervisor checks a small list of critical items to be sure they were not missed by the interviewer or the respondent. If any of these is missing, the supervisor attempts to retrieve the data by calling the respondent. In addition, many studies include a brief reinterview of a random subset of all respondents. The reinterview includes some of the questions in the questionnaire and elicits the re-

TABLE 11-3. *Standard probes used in interviewing*

Repeat question
Repeat part of question (e.g., give the frame of reference again)
• Which would be closer? (repeat answer categories)
• So, would you say that it is? (repeat answer categories)
• Could you be more specific?
• Could you say more about that?
• Are there any other reasons?
• What else?

spond  
tervie

Mar  
that re  
ternat  
the stu  
penhe  
analy  
to enc  
sion l  
Coc  
data e  
errors  
and c  
scann  
Key  
withi  
has be  
swers  
ing, a  
Fin  
group  
oncile  
rates,  
serve

A c  
in th  
meth  
ety a  
recei  
ment  
on w  
Ne  
ologi  
ulatic  
fered  
at co  
field  
as th  
diffe  
St  
divic

respondent's impression of how long the interview took and his or her opinion of the interview.

### DATA PREPARATION

Many questionnaires, abstracts, and other data collection instruments include items that require assignment of codes based on coding schemes either in general use (e.g., International Classification of Disease, Standard Industrial Classification) or devised for the study. Standard practices for coding data are described in detail in other texts (see Oppenheim, 1992). Selection of coding schemes depends on the utility of the scheme in analysis and the ease of use in coding. The coding manual documents the schemes used to encode the data. A fraction of forms are recoded, blindly, if possible. A coding decision log documents decisions not covered by the general rules.

Coded forms are keyed. In some studies, all forms are keyed twice, with the second data entry clerk blind to the already keyed data. This practice ensures against data entry errors but costs more than nonblinded double entry or single entry. At least for identifiers and other critical items, double entry usually warrants the extra expense. Optically scanned forms do not need coding but require care in handling.

Keyed or scanned data are "cleaned" by checking the answers to be sure that they fall within a range of plausible answers (which are overridden after an out-of-range answer has been verified) and that they are mutually consistent ("logic checks"). Inconsistent answers given by the respondent may be flagged, but not changed. Errors in recording, coding, and keying are corrected.

Finally, the numbers of respondents and nonrespondents, overall and within study groups, tracked by the management system during the field phase are compared and reconciled to the numbers in the data file to be used for analysis. The response counts and rates, manuals used during the field phase, study logs, and data collection instruments serve as essential reference documents throughout analysis and presentation of findings.

### CHANGING FIELD METHODS

A decade from now epidemiologists will still be using many of the methods described in this chapter, but many others will have changed. Epidemiologists will devise new methods or adapt those developed for other purposes to epidemiology. Changes in society and technology will impel changes in field methods. Both new and old methods will receive examination in methodologic studies, some focused narrowly, as on measurements of historical exposure to electromagnetic radiation, and others focused broadly, as on ways to improve interviews.

New population sampling methods can be expected to develop. For instance, epidemiologists began using telephone-based sampling for drawing controls from the general population less than two decades ago. In places without lists of residents, these techniques offered an attractive alternative to area-based surveys, which had been used extensively but at considerable cost. Studies comparing control-selection methods will aid in designing field methods for new studies (Lele, 1994). We can expect changes in telephone sampling as the revolution in communications technology continues. In the United States, entirely different sampling frames may emerge as the organization of health care evolves.

Studies of historical cohorts may change dramatically. Laws protecting privacy of individuals may limit possibilities for assembling cohorts. Access to old medical records

and pathology specimens may decline, at least in the United States. On the other hand, computerization of many records may expand greatly the opportunities to assemble cohorts and acquire outcome data.

Increasing numbers of epidemiologic studies will incorporate genetic and other biologic markers. Associated methodologic studies are needed to understand the results of each assay and compare assays in various ways. Conduct of a biochemical epidemiology study requires answering a myriad of logistical questions: Do the assays require a sample of serum or can urine, breath, hair, etc., be used; how compliant will subjects be; would one designated staff member collect the samples more reliably than several; how long can the material be kept before shipping without deterioration; are split samples necessary; and so forth? The data to guide these decisions may be lacking, but studies to compare variations in procedure often can be incorporated into the main study.

Finally, many widely used methods still need to be studied in particular applications (e.g., questionnaires to assess past diet or occupational exposures). The performance characteristics and costs of many field methods are unknown. Many studies are needed to measure the impact of various field methods on likely levels of bias. Even general aspects of quality control need evaluation, for example, audiotaping of questionnaires for interviewer performance (Edwards et al., 1994). Such quality control techniques themselves deserve examination, in this instance, a test of whether the presence of the tape recorder influences the answers.

Means to reduce nonresponse command attention. Such studies often can be embedded in the main study, as in an investigation of the impact of leaving messages on telephone-answering machines (Harlow et al., 1993). In a study with a mail questionnaire, investigators included a randomized trial of asking refusers to explain their refusal and found the response rate was virtually unaffected (Shahar et al., 1993). Assessment of the impact of nonresponse also needs continued study. For instance, a comparison of early and late respondents to postal surveys showed little difference in a variety of health indices (Paganini-Hill et al., 1993). These data aid in deciding how much field effort to devote to increasing the response rates.

Misclassification of exposure continues to plague epidemiology, so studies of validity and reproducibility of exposure assessment abound. Continuing studies of data from proxy respondents are needed (Byers et al., 1993; Johnson et al., 1993), as are studies of recollection (Gilpin and Pierce, 1994; Norrish et al., 1994).

Epidemiologic design and analysis have changed greatly in recent years, and epidemiologic field methodology has changed at least as much. The continuing investigation of field methods promotes incremental improvement and exploration of radically different techniques. Improvement in field methods offers the continuing promise of more powerful epidemiologic studies that reveal more of the causes of human disease.

#### ACKNOWLEDGMENTS

We are grateful to Louise Brinton and Marcia Russell for contributing their unpublished data. We also would like to thank Shelia H. Zahm, Michelle Doody, Patricia Stewart, Ruth Kleinerman, Mary Fraser, Sholom Wacholder, and Linda Lannom for providing valuable advice.