requires some period of "regular" smoking for an individual to be classified as an ever smoker. 128 of 252 individuals reported being never smokers. However, when assessed concurrently with another questionnaire in which regular smoking was not defined and the respondent self-defined smoking, 7 percent fewer subjects (119 of 252) reported being never smokers.

Thus, the use of more clearly defined questions, such as specifying 100 cigarettes in a lifetime, or 1 cigarette per day for 1 year, or 5 cigarettes per week for 1 year, will reduce misclassification. However, some misclassification will still occur for those individuals who smoked for relatively brief periods during their lives but cannot accurately remember how long they smoked or accurately estimate the number of cigarettes they smoked.

Attention also must be paid to defining current or former smokers. Some studies, such as the Cancer Prevention Study I (CPS-I) (Hammond and Garfinkel 1969), define current smokers as those who respond affirmatively to the question "Have you smoked within the past year?" Other studies use smoking in the past 6 months as the guideline for current smokers (Coultas et al. 1988). The criteria for questions identifying current smokers can range from having smoked in the past year, to the past 6 months, to the past week, or to an unspecified period. A few additional questions will enhance the specificity of the definitions of current smokers and former smokers. These items, or comparable ones, have been used in previous surveys, for example, the 1988 Baseline Prevalence Survey for the Community Intervention Trial for Smoking Cessation, funded by the National Cancer Institute: "At what age did you start smoking on a regular basis?"; "On the average, about how many cigarettes did you smoke per day during the last 12 months you smoked?"; and for former smokers, "When did you quit smoking cigarettes?" (recorded to exact date if possible). These items provide additional information for defining ever smokers, or stratifying by levels of exposure, and for determining the period of abstinence.

The dynamic nature of smoking cessation highlights the importance of being aware that any categorical definition of former smoker in relation to the health effects of smoking cessation will include former smokers who have been abstinent for varying periods of time. Optimally, questions on smoking history should ascertain the duration of abstinence for former smokers, and if possible, abstinence periods should be treated as continuous or categorized variables in an analysis, thus avoiding the problem of treating former smokers as a single group. However, benefits of cessation are still clearly observed in spite of the limitations of using categorical data.

The most common minimum periods of abstinence used for defining former smoking status are 24 hours, 7 days, and 30 days. The National Interagency Council on Smoking and Health (1974) recommended using a minimum of 7 days of abstinence for defining cessation. However, because of the nature of smoking, using a short abstinence period to define former smokers is not optimal in epidemiologic studies. The degree of misclassification of former smokers will depend on the minimum duration of abstinence used to define former smokers and the criterion used to consider determine relapse.

Many studies do not specify a minimum duration of abstinence for individuals classified as former smokers at a particular point in time. Data from such studies on the association of smoking cessation with health and disease outcomes must be

interpreted cautiously. For example, in the reports of the Whitehall Civil Servants Study (Rose and Hamilton 1978; Rose et al. 1982), the criterion used to define abstinence is not indicated. The only information provided is that the smokers reported that "they were then smoking no cigarettes at all" (Rose and Hamilton 1978).

Regardless of the criteria used to define abstinence, the methodology for assessing smoking status, including questionnaire items, needs to be carefully described by investigators. Optimally these items should enhance the process of obtaining information regarding the duration of abstinence, making it possible to fully determine the relationship of smoking cessation to health and disease outcomes. When reviewing studies of the health effects f smoking, the definition of the former smoker must be carefully assessed, and the effect of the definition on the findings must be carefully examined.

Temporal and Frequency Issues

Studies vary according to whether smoking is assessed retrospectively or prospectively and whether a single assessment or a series of assessments is used. The category of never smokers can be assessed retrospectively, usually relying on a single assessment. Requiring subjects to reconstruct more detailed smoking histories can be very demanding. Nevertheless, simply classifying individuals as former smokers or current sm[illegible]s reveals very little about the amount of smoking exposure experienced. More pert[illegible]ent questions regarding exposure include "How long have you been abstinent from cigarettes?"; "At what age did you start smoking?"; "How many cigarettes did you smoke during different periods of your life?"; "How many times did you stop smoking?"; and "How long did you remain abstinent during each of these occasions?"

A series of repeated assessments can result in inconsistencies such as some individuals reporting smoking at one assessment and later reporting that they never smoked. In a followup study in England, for example, Britten (1988) found 1,296 participants aged 36 who claimed that they had never smoked. Of these, 242 (18.7 percent) previously had reported smoking less than 1 cigarette per day, and 102 (7.9 percent) previously had reported smoking at least 1 cigarette per day for at least 1 year. Of the 102 who reported previously that they had been regular smokers, 93 percent reported that the last time they had smoked was at least 10 years prior to the survey.

If the Britten study had used only one retrospective assessment of the subjects at age 36, 32.5 percent of the 1,296 subjects would have been classified as never smokers and 32.6 percent as former smokers. Assuming that reports at a young age were more accurate because memory bias was less likely to occur, the serial assessment indicates that a more accurate categorization would be 29.1 percent for never smokers and 36.5 percent for former smokers. Britten (1988) estimated that misclassification of this magnitude, when applied to a study by Friedman and colleagues (1979), would result in only a 5-percent increase from 2.41 to 2.53 in relative risks of death for former smokers compared with never smokers.

Krall and colleagues (1989) found that of 87 middle-aged adults, 87 percent accurately recalled their smoking status of 20 years earlier, but only 71 percent accurately recalled the amount that they had smoked. Furthermore, underestimation of the amount

smoked was twice as common for 20 years earlier (17 vs. 9 percent) and six times more common for 32 years previously (37 vs. 6 percent). Persson and Norell (1989) found that in a random sample of 9,394 individuals in Sweden, retrospective information obtained 6 years later resulted in a strong tendency to overestimate previous cigarette consumption among individuals who had increased their smoking (69 percent overestimated) and to underestimate among individuals who had decreased their smoking (49 percent underestimated). Subjects with unchanged cigarette consumption showed the highest levels of agreement (89 percent) between original and retrospective information. Rather than reconstructing full smoking cessation histories that are subject to biased reporting, many retrospective studies rely on more limited categorization such as never, former, and current smokers.

Retrospective studies enable researchers to assess long periods of smoking abstinence without the need to observe the subjects over a long period of time, as would be necessary in prospective studies. Case–control studies, for example, can compare cases with smoking-related diseases with controls with histories of being abstinent for 10 to 20 years; in a prospective study, it may be impractical or impossible to study health consequences of cessation with more than 10 to 20 years of abstinence (Chapter 2, Part II).

Prospective studies have the potential for more reliable and valid measures of smoking status over time, especially when using a series of assessments, than do retrospective studies. In intervention trials, for example, all subjects enter the trial as current smokers. Following intensive intervention, subjects are identified as continuing smokers or former smokers (abstinent). By assessing subjects at specified intervals such as every 4 or 6 months over a series of years, especially when paired with biochemical verification (Chapter 2, see section on Biochemical Markers), researchers can reduce the measurement bias and be more confident in the reliability and validity of measures classifying continuing and former smokers and specifying length of abstinence for former smokers. In MRFIT (Ockene et al. 1990) for example, a series of 4-month followups over 6 years enabled researchers to classify participants into three categories: persistent quitters (continuous abstainers since the initial intervention), intermittent quitters (abstinent for periods of time since the initial intervention), and continuous smokers (not abstinent during any of the followup periods). Such precision in measurement is generally not possible or necessary in epidemiologic studies.

Prospective studies may use a single assessment to categorize current, former, and never smokers. These studies then prospectively examine the categories to detect differential rates of morbidity and mortality. As discussed above, the assumption that individuals will not change their smoking status maybe a flaw with such single assessments.

Improving Self-Report Measures

Ideally, assessments of smoking status need to include standardized questions to determine smoking status, that is never, current, and former smokers. For example, to be categorized as a never smoker, the necessary response would be "no" to a standard question such as, "Have you ever smoked at least 1 cigarette per day for at least 1 year?"

Whenever possible, questions should be used that allow continuous rather than dichotomous scales for response. A question such as "Do you smoke regularly?" results in a dichotomous response scale. This scale provides much less information than does a continuous scale, such as the question, "On the average, how many cigarettes do you smoke per day?" which can range from 0 to 20, 40, 60, or more. Multiple questions such as, "Have you smoked even a puff of a cigarette in the past 7 days?"; "How many cigarettes do you typically smoke each day?"; and "How many cigarettes do you typically smoke each week?" can be used to refine a category such as current smokers. Inclusion of other indices, such as biochemical markers of smoking (e.g., saliva cotinine levels), can also be used to describe smoking status.

In a followup study, measures of smoking status optimally should be repeated over multiple occasions, especially for dynamic categories like current smokers and former smokers, which are open to change over time. Repeated measures over a series of occasions provide further reliability and validity for assessments and also provide greater statistical power for detecting differences between groups. Nevertheless, studies with only a single or a few assessments of smoking behavior have been extremely informative.

Alternative Behavioral Measures

As a measure of smoking, self-report by questionnaires and interviews is the most common, the least expensive, the easiest to use, and the most feasible in epidemiologic studies (Frederiksen, Martin, Webster 1979; Pechacek, Fox et al. 1984). However, other behavioral measures have also been used in clinical studies. Because these measures are generally not used in large-scale epidemiologic studies, they will be presented only briefly in this Chapter.

Self-monitoring by the smoker, a measure of smoking commonly used in intervention studies, involves recording by paper, pencil, and mechanical counters each cigarette as it is smoked. The monitoring itself may be a reactive measure and alter the behavior, depending on the nature of the monitored behavior and motivation (Abrams and Wilson 1979; Frederiksen, Martin, Webster 1979; Lipinski et al. 1973; McFall 1978; Orleans and Shipley 1982). It is an intrusive measure that is normally restricted to small studies of high intensity. Other behavioral measures, such as direct observation, collecting and counting cigarette butts (McFall 1978), and measuring their length (Auger, Wright, Simpson 1972), are even more costly and intrusive and less appropriate for epidemiologic and large intervention studies.

Alternative types of behavioral reports for validation of smoking status include verification by an informant (Shipley 1981), by self-report measures using multiple questions about smoking behavior or status as part of the same interview or questionnaire (see above), and by sampling on multiple occasions. Examples of the latter usually involve long periods of time and often result in multiple sources of discrepancy. (See Lee 1988 for summary.)

Surrogate Assessments

In some circumstances researchers may need to obtain information from sources other than the index subjects. With some study designs, for example a case–control study of lung cancer, some subjects are unavailable to answer questions because of illness or death. In cohort studies, or intervention studies with mortality endpoints, surrogate interviews are sometimes required to assess smoking during the interval preceding death.

Failure to obtain surrogate reports can cause considerable bias in some instances. In a case–control study of oral cancer, Greenberg and coworkers (1986) obtained interviews with 112 cases (67.9 percent) and surrogate reports for 23 cases (13.9 percent). Cases needing surrogate reports had more advanced stages of disease at the time of diagnosis and were more likely to be black and less educated than cases interviewed in person. Cigarette smoking and drinking hard liquor were more common among these cases. Therefore, failure to include surrogate reports would have resulted in underestimates of the strength of association between cigarette exposure and hard liquor and the risk of oropharyngeal cancer.

Pickle, Brown, and Blot (1983) found that siblings of index subjects provided the most complete data about smoking in the subject's family of origin and early life events. Spouses and offspring supplied the most complete data about smoking history during adult life. Incomplete data generally increased with the amount of detail requested, so that there were considerably higher nonresponse rates for a detailed smoking history (approximately 50 percent) than for the history of a broad smoking status, such as never smoker (approximately 15 percent). Surrogates beyond a spouse or close relative provided much higher nonresponse rates for almost all questions in all statuses.

McLaughlin and colleagues (1987) examined the reliability of retrospective surrogate reports obtained 10 years after initial reports and compared these with retrospective self-reports using data from the NHANES-I (Cornoni-Huntley et al. 1983). Correct identification of previous smoking status was generally provided by most types of surrogates, except siblings of male decedents. The combined level of agreement for all surrogates ranged from 85 to 95 percent and was remarkably similar to that from self-reports of living subjects. Thirty-five percent of the surrogates could not provide data on when smoking began compared with 1 percent in self-reports. Surrogates who responded tended to provide a later age for starting. Surrogates did, however, provide estimates of years smoked that were comparable to the original reports. In this study, siblings and other surrogates provided less reliable reports than spouses, offspring, or parents of subjects.

Lerchen and Samet (1986) interviewed widows of lung cancer patients who had supplied their own smoking histories while alive. They found that of 77 wives of current smokers, all supplied information about the cases' cigarette smoking status (ever/never) that was in perfect agreement with the information supplied by the cases themselves. Sixty-six (86 percent) were able to supply complete responses about their husbands' smoking behavior. For those who responded, however, mean values reported by cases and their wives were not significantly different for age at which cases started smoking, years smoked, or average number of cigarettes smoked per day. Wives tended to report

20 cigarettes smoked daily even when their husbands smoked substantially more or less. Pershagen and Axelson (1982) also reported perfect agreement regarding smoker/nonsmoker status when information was obtained from a close relative (parent, wife, or child) for 14 lung cancer cases compared with information that had previously been obtained from the cases by the physician. Blot, Akiba, and Kato (1984) also interviewed next of kin in a case–control study of lung cancer among atomic bomb survivors who had previously provided information regarding their own smoking behavior while they were alive. The investigators found that only 1 percent of surrogates reported that a subject had never been a smoker while the subject reported that he or she had smoked, suggesting that the identification of never smokers by next of kin is very accurate. There was poorer agreement regarding those who smoked, with 13 percent of surrogates indicating that a subject had smoked while the subject had reported never smoking.

Sandler and Shore (1986) examined the quality of data provided by adult offspring on parents' smoking and drinking. The data were from 518 cancer cases and 518 healthy controls aged 15 to 59. When possible, mothers provided data on their own smoking and their husbands' smoking. Of 982 subjects who had lived with their natural mother, 97 percent provided data on their mothers' smoking status. Of those whose mothers reported never having smoked cigarettes, 2.7 percent were reported as ever having smoked by the adult child. Of those mothers who reported ever having smoked, 8.8 percent were reported as never smokers. Of those fathers reported by the mother as never smokers, 17.2 percent were reported by subjects as ever smokers. Of those fathers reported as ever having smoked cigarettes, 21.1 percent were reported as never smokers by their adult children. Even with the quantity of cigarettes collapsed into categories to include answers of less than 1 pack, 1 pack, and more than 1 pack, the proportion of mothers and subjects whose responses exactly agreed was 82.0 percent for mothers and 49.2 percent for fathers.

Humble, Samet, and Skipper (1984) interviewed 46 subject–spouse pairs, with 2 people in each of 38 of these pairs acting as the subject and as a surrogate for his/her spouse, thus producing 84 total subject–surrogate pairs. For the 30 current or previous cigarette smokers whose spouses gave complete smoking data regarding the subjects, the subjects reported a mean use of 17.8 cigarettes per day compared with 14.3 reported by their spouses. The difference was not significant.

Investigations indicate that useful information on smoking can be obtained in epidemiologic investigations that must rely on surrogate information (McLaughlin et al. 1987). Although greater misclassification occurs when surrogate reports are used compared with self-reports, consideration of variables such as the relationship of the informant, length of time he or she had known the case, the topic of the questions, and complexity of the data gathered from the informant can add to the validity of the data (Rogot and Reid 1975).

Nonbehavioral Measures

Methods other than self-report have been used to assess smoking status. Some researchers have expressed concern that self-report when used alone can be an in-

accurate measure that underestimates the amount of cigarettes smoked (Haley and Hoffmann 1985; Marsh et al. 1988; Warner 1978) because subjects often underreport levels of cigarette consumption or misrepresent themselves as former smokers (Luepker et al. 1989; Murray and Perry 1987; Windsor and Orleans 1986; Russell 1982; Stookey et al. 1987). Underreporting also has been linked to "digit bias," that is, subjects tend to report in terms of multiples of ten and underestimate actual consumption (Pechacek, Fox et al. 1984; Vogt 1977; US DHHS 1989).

Between 1974 and 1985, estimates of U.S. cigarette consumption based on self-report accounted for only about 70 percent of consumption estimates based on cigarettes taxed and sold (Hatziandreu et al. 1989). This ratio has remained relatively stable. Most of this discrepancy is likely to be due to underreporting or a "rounding down" to the nearest multiple of a half-pack of daily cigarette consumption (Kozlowski 1986), although misreporting of smoking status may play a role as well.

Validation of self-reports with measures such as biochemical assessments represents a possible means of decreasing misclassification due to misreporting (Luepker et al. 1989; Windsor and Orleans 1986). However, some researchers note that biochemical validation techniques present different problems that also cause misclassification, thus favoring the use of self-report (Assaf et al. 1989; Crossen, Dougher, Belew 1984; Hansen, Malotte, Fielding 1985; Hatziandreu et al. 1989; Kornitzer et al. 1983; Petitti, Friedman, Kahn 1981). As noted above, sensitivity and specificity of the biochemical measures are not perfect. In addition, the procurement of biochemical measures from a large majority of self-reported quitters is not as feasible in large-scale intervention trials or observational studies as it is in smoking studies of a smaller scale and a more clinical nature. Subjects in the population samples do not have the same commitment to studies that volunteers have to clinical studies, and the former are more likely to leave the study area, which makes validation difficult (Ockene et al. 1989). Validation also requires more personal contact than is generally employed in observational or large-scale field studies, and the additional contact may not be acceptable to the subjects or feasible in the context of the study.

The section below on physiologic measures discusses methods other than behavioral measures that have been used to assess cigarette smoke exposure. These measures are then contrasted with self-report, and the varying needs for biochemical measurement among different populations are considered.

Physiologic Measures

Smoking behavior has been assessed by measuring physiologic changes that result from smoking (Pechacek, Fox et al. 1984). Smoking and smoke exposure are reflected in a variety of acute and chronic physiologic measures primarily because of the strong pharmacologic effects of nicotine. These effects include changes in heart rate, blood pressure, hand tremor, and skin temperature. Each of these measures has a wide variability under normal conditions and is affected by many factors other than smoking, thus limiting usefulness as a measure of smoking (Pechacek, Fox et al. 1984).

Biochemical Markers

Cigarette smoke is a complex mixture of chemicals, some of which are present in the tobacco leaf and some of which result from chemical reactions during either the curing process or smoking (US DHEW 1979; US DHHS 1986, 1989). Three chemical constituents of tobacco smoke, carbon monoxide (CO), hydrogen cyanide (HCN), and nicotine, pass through cigarette filters and are present in inhaled tobacco smoke in concentrations high enough to be absorbed and detected in persons who smoke. These chemicals are measurable as intact compounds or as metabolic products.

Exposure to CO can be assessed in the blood as carboxyhemoglobin (COHb) or as CO in expired alveolar air. Methods are available for measuring cotinine, the primary metabolite of nicotine, and SCN^-, a metabolite of HCN, in urine, blood, and saliva. Other measures, such as skin-surface sampling for nicotine (Nanji and Lawrence 1988) are not as well established.

Extensive reviews of the literature on the use of biochemical markers as measures of smoking status are provided by Benowitz (1983), Haley and colleagues (1986), Lee (1988), Pechacek, Fox, and colleagues (1984), and Windsor and Orleans (1986). Cummings and Richard (1988) supplied a review of optimal cutoffs for the biochemical measures discussed here. This Section is not intended to provide an indepth review of the variability and biochemical rationale for these measures and will only provide an overview of the use of biochemical assessments for smoking status.

Terminology

Sensitivity and specificity, characteristics of a test such as a biochemical assessment, are measures of validity, the extent to which the test measures truth (Fletcher, Fletcher, Wagner 1987). Typically, sensitivity and specificity are determined by comparing the test results against a reference or "gold" standard. For smoking, self-reported status has most often been used as the standard for assessing biochemical markers. The sensitivity of a biochemical test for smoking exposure is the proportion of true smokers who are classified as smokers by the biochemical test. The specificity of a biochemical test for smoking exposure is the proportion of true nonsmokers who are classified as nonsmokers by the biochemical test. A test of 100-percent sensitivity and 100-percent specificity would perfectly discriminate true smokers from true nonsmokers. However, this degree of validity is not reached by any presently available biochemical marker. In addition, the standard to which biochemical measures are compared, typically self-reported smoking status, may be of limited validity, and thereby cause apparent sensitivity and specificity to be reduced.

When continuous measures are used to test for smoking status, a cutpoint must be chosen such that those individuals whose test value exceeds the cutpoint are classified as smokers and those with values below the cutpoint are classified as nonsmokers (Cummings and Richard 1988). The level at which the cutpoint is set determines the sensitivity and specificity of the test. Lowering the cutpoint improves the sensitivity at the expense of specificity. Raising it will improve specificity at the expense of sensitivity (Cole and Morrison 1980; Browner, Newman, Cummings 1988). Selecting

a cutpoint depends on the relative importance of mislabeling an actual smoker as a nonsmoker with a very insensitive but specific test versus mislabeling an actual nonsmoker as a smoker with a very sensitive but nonspecific test. This tradeoff between sensitivity and specificity is discussed in more detail elsewhere (Fletcher, Fletcher, Wagner 1987).

An important contextual issue concerns the validity with which the biochemical measure classifies individuals. When the test is applied to a population of smokers and nonsmokers, the proportion of the persons who test positive, that is, above the specified cutpoint, who are actually smokers becomes an important concern. This issue, distinct from the question of what proportion of smokers are above the cutpoint, is the crucial measure of how much misclassification occurs. This proportion, the positive predictive value of a test, depends not only on specificity and sensitivity but also on the prevalence of the condition in the population being tested (smoking in this example). The less prevalent smoking is in the screened population the lower the positive predictive value of a test (Browner, Newman, Cummings 1988).

The relative misclassification rates for smokers and nonsmokers, determined in part by the estimated prevalence of smoking in the population to which the cutpoints are applied, are particularly important in studies which use biochemical tests to verify self-reported smoking cessation (Cummings and Richard 1988; Ruth and Neaton, in press). For example, the pressure to quit smoking that is present in formal smoking cessation programs may result in a high proportion of continuing smokers who report not smoking. The use of cotinine validation in such circumstances (high prevalence of false reporting) results in a high positive predictive value, as opposed to the lower positive predictive value when the same test is applied to self-reported former smokers identified in a population-based survey (low prevalence of false reporting).

In biochemical validation studies, such as those reported in a subsequent section of this Chapter, after optimal cutpoints are set using self-report in one population as the gold standard, the biochemical marker then becomes the gold standard against which self-reported smoking status is measured in another population.

Carbon Monoxide

High concentrations of CO are present in cigarette smoke (US DHEW 1979; US DHHS 1986, 1989). Absorbed rapidly into the bloodstream during smoke inhalation, CO has a half-life of 4 to 5 hours in sedentary adults (Stewart 1975). Direct measurements of CO can be taken from exhaled alveolar air or estimated by measuring the percentage of hemoglobin combined with CO (COHb) (Stewart 1975).

Sensitivity of exhaled CO for classifying active smoking is generally in the range of 80 to 85 percent but can be affected by diurnal variability as well as other factors (Benowitz 1983). Given the short half-life of CO, levels are influenced by time of day and time elapsed since last cigarette. Measurements taken late in the day, standardized from time since last cigarette, are likely to give the best estimates of CO levels (Frederiksen and Martin 1979; Horan, Hackett, Linberg 1978; Hughes, Frederiksen, Frazier 1976). Using self-report of recency of smoking can increase sensitivity (Bauman, Koch, Bryan 1982). Sensitivity is poor for light smokers (Fortmann et al.

1984; Vogt 1982), and specificity can be reduced by exposure to CO present in the environment as a result of industrial and automobile pollution, environmental tobacco smoke, indoor combustion sources, and use of products such as marijuana (Biglan et al. 1985; Frederiksen and Martin 1979; Stewart 1975). In spite of this, only 2 to 5 percent of nonsmokers in general populations will exceed 1 percent COHb (Janzon et al. 1981; Kahn et al. 1974). Using COHb levels from a national probability sample, the Radford and Drizd (1982) reported the 95th percentile for COHb to be 1.77 percent in nonsmokers, aged 12 to 74. If a 2-percent cutpoint is applied to this sample, 3.6 percent of nonsmokers would be incorrectly classified as smokers.

Thiocyanate

High concentrations of HCN, a toxic gas, are present in cigarette smoke. However, HCN is very active chemically and is rapidly detoxified by the liver into SCN^- (Langer and Greer 1977; Boxer and Rickards 1952). Because SCN^- accumulates in body fluids, such as saliva, urine, and blood, it is used as a biochemical measure of exposure to tobacco smoke. The biologic half-life of SCN^- has been found to vary quite a bit (Bliss and O'Connell 1984) although the length of time usually noted is between 10 and 14 days (Langer and Greer 1977; Vesey 1981). Salivary SCN^- can be measured most reliably in parotid gland secretions (Shannon, Suddick, Dowd 1974); however, parotid gland secretions show some seasonal and diurnal variability (Shannon, Suddick, Dowd 1974). When serum and saliva samples are compared, the levels are 15 to 20 times higher in saliva than serum (Langer and Greer 1977; Pechacek et al. 1979; Vesey 1981). However, saliva levels are more variable (Pechacek et al. 1979).

The increment of SCN^- in light smokers is low, and there is much overlap of SCN^- levels in light smokers compared with nonsmokers (Fortmann et al. 1984; Neaton et al. 1981; Vesey et al. 1981). However, detection of light smoking in adults using SCN^- levels is better than in adolescents (Windsor et al. 1985). This is likely to be related to the fact that adolescents are often in the process of learning how to smoke and inhale, and they may not have an established pattern of smoking (Pechacek, Murray et al. 1984). For example, among younger adolescents only one-third or less could be identified on a single assessment (Hunter, Webber, Berenson 1980; Luepker et al. 1989; Pechacek, Murray et al. 1984). Specificity represents a more severe problem than sensitivity. A large number of food products are sources of either cyanogenic glycosides (e.g., almonds, bamboo shoots, sugar cane) or naturally occurring SCN^- (e.g., cauliflower, broccoli, beer) and can produce levels of SCN^- in saliva equivalent to the average levels of smokers (Langer and Greer 1977; Neaton et al. 1981; Pechacek et al. 1979; Swan et al. 1985).

The relatively low specificity and sensitivity of SCN^- testing compared with cotinine and CO make SCN^- a less useful outcome measure for smoking cessation studies (Gillies et al. 1982; Fortmann et al. 1984) unless adjustments are made using carefully collected dietary and environmental exposure data. A prime advantage of using SCN^- for biochemical validation of smoking abstinence is its long half-life compared with other biochemical measures (Fortmann et al. 1984; Steinman 1985; Murray et al. 1987;

Pechacek, Fox et al. 1984), which is of particular interest in population surveys where longer term abstinence is of concern.

Cotinine

Cotinine, a metabolic byproduct of nicotine, is distributed throughout extracellular fluid and is excreted through the kidneys and salivary glands (Benowitz 1983). About 15 to 20 percent is eliminated in the urine unchanged, and the rest is metabolized (Benowitz 1983). The half-life estimates of cotinine are variable and range from 15 to 40 hours (Carey and Abrams 1988; Knight et al. 1985; Greenberg et al. 1984; Haley and Hoffmann 1985; Haley et al. 1987; Sepkovic, Haley, Hoffmann 1986). The differences in estimated half-life for cotinine reflect not only individual differences in metabolism but also differences between smokers and nonsmokers (Haley, Sepkovic, Hoffmann 1989; Sepkovic, Haley, Hoffmann 1986; Haley et al. 1987). Cotinine levels vary with the diurnal cycle and are best assessed late in the day (Benowitz 1983). Methods are available for measuring cotinine in saliva, urine, and blood. Urinary levels have been suggested to be too variable (Pechacek, Fox et al. 1984), and plasma or serum levels appear to be the most stable (Benowitz 1983). However, sampling saliva because of ease of procurement and accuracy in classifying smokers and nonsmokers has been recommended as a useful, noninvasive method that can be applied to large-scale intervention trials (Abrams et al. 1987).

Because nicotine is unique to tobacco, cotinine is a highly valid marker for almost any tobacco use (Haley, Axelrad, Tilton 1983; Russell et al. 1981; Wald et al. 1984; Zeidenberg et al. 1977). Although nicotine has been assessed in some studies, it is recommended that cotinine be used because it has a more enduring and stable blood level (Langone, Gjika, Van Vunakis 1973). Detecting regular smokers by analysis of cotinine in blood, urine, or saliva is almost certain, and even light smokers and intermittent smokers are easily detected (Benowitz 1983; Haley, Axelrad, Tilton 1983; Paxton and Bernacca 1979; Zeidenberg et al. 1977; Carey and Abrams 1988; Williams et al. 1979). In one investigation, 95 percent of adolescent ever smokers were detected by cotinine (Williams et al. 1979). Specificity is also high; regular smokers typically have blood cotinine levels of 200 to 400 ng/mL, light smokers have 40 to 50 ng/mL, and nonsmokers are typically below 10 ng/mL. When nonsmokers are assessed, they rarely have any detectable cotinine (Benowitz 1983; Haley, Axelrad, Tilton 1983; Sepkovic and Haley 1985; Zeidenberg et al. 1977).

In comparative studies of different biochemical measures of smoking, cotinine has emerged as the measure of choice (Abrams et al. 1987; Haley, Axelrad, Tilton 1983; Jarvis et al. 1984, 1987; Knight et al. 1985; Pojer et al. 1984) because of its superior sensitivity and specificity. However, it is more expensive and more analytically complex than the other biochemical measures.

The value of biochemical measures is limited to short-term abstinence and cannot be used to document continuous abstinence in long-term studies. CO, with a half-life of 4 to 5 hours, can validate self-reports of not having smoked in the past 24 to 48 hours (Benowitz 1983). Cotinine, with a half-life of 15 to 40 hours, would have limited application for validation beyond a few days. SCN^-, with a half-life of 10 to 14 days,

has been used to validate self-reports of not having smoked in the past 7 days and may be useful to validate up to 3 to 4 weeks. However, specificity of this measure is low compared with cotinine and CO.

Bogus Pipeline

The bogus pipeline, an assertion to subjects that biochemical assessments will be used to assess smoking status when they will actually only be collected but not evaluated, is used mostly in research with adolescents. One of the reasons given by researchers for continuing to use biochemical verification for at least some proportion of the total subjects is the assertion that if the subjects believe biochemical validation will occur, they will be more likely to provide valid responses to self-report measures. This "bogus pipeline effect" was first presented by Evans, Hansen, and Mittelmark (1977) from the work of Jones and Sigall (1971) concerning smoking among adolescents. It is believed that there is great pressure among adolescents to misreport smoking activities. Murray and coworkers (1987) provided an extensive review of this aspect.

Murray and Perry (1987) attempted to determine the conditions under which a bogus pipeline will be effective by manipulating conditions of anonymity. They demonstrated that a bogus pipeline for adolescents is more likely to have an effect if there is an expectation that subjects would otherwise perceive large amounts of pressure to report not smoking and there is a credible pipeline message. However, their findings suggest that an effective procedure to ensure anonymity can reduce this pressure and likewise reduce the need for the pipeline.

Contextual Issues Affecting Biochemical Assessment

The accuracy of self-report measures, the desirability for behavioral or biochemical validation of self-report, and the type of assessment needed are issues that need to be considered in the context of the type of study, the nature and size of the study sample, and possible refusal problems.

The nature of the subject sample can affect the likelihood of misreporting and therefore the desirability of validation by biochemical assessment. In Table 1, studies demonstrating misreporting rates for individuals who report cessation but who are assessed to be smokers by cotinine or nicotine measurement are classified into three types of subjects: untreated volunteer samples, intervention samples, and high-risk for disease and/or medical patients. Table 2 presents a similar classification of studies demonstrating misreporting with CO validation. The tables are adapted from Lee's work (1988) with the inclusion of additional studies. In cases where multiple cutoff criteria are recorded, the values closest to the optimal cutoff are reported. Several studies should be viewed as outliers and are noted in the tables. These studies reported unusually high rates of individuals who reported not smoking but were above the cutpoint and also employed cutoff criteria far below optimum cutpoints (Cummings and Richard 1988).

For untreated volunteer samples, the mode for individuals classified as smokers by biochemical assessment who reported not smoking is zero, and no sample exceeds 5

TABLE 1.—Measures of false reports of not smoking from studies using nicotine and cotinine as a marker

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Part I. Volunteer samples					
Russell and Feyerabend (1975)	London smokers, nonsmokers, and heavy passive smokers	No	Urinary nicotine	0 (0/27)	No overlap between range of urinary nicotine levels of smokers (N=18) and nonsmokers (N=27)
Williams et al. (1979)	Students health screening	No	Plasma cotinine	2 (2/98)	
Haley, Axelrad, Tilton (1983)	New York nonsmoking volunteers	No	Salivary or plasma cotinine	0 (0/18)	No cutpoint established; no cotinine detected in nonsmokers
Wald et al. (1984)	Nonsmokers attending BUPA[a], and Oxford colleagues	No	256 ng/mL urinary cotinine	0.9 (2/221)	Cutpoint based on distribution
Haddow, Palomaki, Knight (1986)	US women attending well-women screening	No	30 ng/mL serum cotinine 10 ng/mL serum cotinine	1.3 (3/232) 2.2 (5/232)	
Coultas et al. (1987)	New Mexico Hispanic children and adults in household survey	No	50 ng/mL salivary cotinine	3.2 (43/1,360)	46.3% of sample below age 18 yr
Lee (1987)	Representative UK sample providing saliva, without prior warning, after smoking data	No	30 ng/mL salivary cotinine 10 ng/mL salivary cotinine	2.5 (20/808) 4.2 (34/808)	
Nanji and Lawrence (1988)	Lab sample	No	1 μg/mL skin nicotine	0 (0/43)	
Pierce et al. (1987)	Sydney, Melbourne smokers	No	250 nmol/L salivary cotinine	4.0 (25/622)	

TABLE I.—Continued

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Part II. Intervention samples					
Russell et al. (1979)	London smokers attending general practices in intervention trial	Yes	Salivary nicotine	7.1 (1/14)	No cutpoint established; length of followup not stated
Paxton (1980)	UK smokers assigned to various stop treatments	Yes	Urinary nicotine	n=2, N<60	Study began with 60 subjects; 2 false reports of not smoking detected; cutpoint not established; 6-mo followup
Jamrozik, Vessey et al. (1984)[b]	UK smokers attending general practitioners in trial of various antismoking interventions	Some groups	100 ng/mL urinary cotinine	23.9 (11/46)	If nonparticipants considered as false reports of not smoking, then 39.7% (23/58) gave erroneous reports; 1-yr followup
Russell et al. (1987)[b]	UK smokers attending general practitioners in trial of effects of brief intervention and support of a smokers' clinic	Some groups	50 µg/L urinary cotinine	38.8 (57/147)	1-yr followup
Abrams et al. (1987)	Smokers/nonsmokers in worksite cessation program	Yes	10 ng/mL salivary cotinine	9.1 (1/11)	Self-reported abstainers; 8-wk followup
Stookey et al. (1987)	Cessation study	Yes	10 ng/mL salivary cotinine	Nonsmokers 0 (0/20) Former smokers 45.1 (46/102)	Length of followup not stated

TABLE 1.—Continued

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Part III. High-risk/medical patients					
Wilcox, Hughes, Roland (1979)	Nottingham MI patients	Yes	2 μg/100 mL urinary nicotine or 10 μg/100 mL urinary cotinine	16.3 (8/49)	An additional 5 subjects had detectable levels in concentrations below the cutpoint
Jarvis et al. (1987)	Clinic outpatients	No	13.7 ng/mL serum cotinine	19 (23/121)	
		No	14.2 ng/mL salivary cotinine	18 (22/121)	
		No	49.7 ng/mL urinary cotinine	17 (21/121)	
		No	21.8 ng/mL salivary nicotine	14 (17/121)	
		No	2.3 ng/mL plasma nicotine	14 (17/121)	
		No	58.6 ng/mL urinary nicotine	16 (19/121)	
Haddow et al. (1987)	US pregnant women	No	10 ng/mL serum	4.9 (142/2,871)	Unpublished data

NOTE: n/N=number of individuals reporting not smoking but with levels of biochemical marker exceeding cutpoint divided by all individuals reporting not smoking; MI=myocardial infarction.

[a]British United Providence Association Medical Center in London.

[b]Studies classified as outliers due to low criterion cutoffs.

SOURCE: Adapted from Lee (1988).

TABLE 2.—Measures of false reports from studies using CO as a marker

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Part I. Volunteer samples					
Jones, Commins, Cernik(1972)	London taxi drivers	No	6.6% COHb	4.8 (1/21)	
Petitti, Friedman, Kahn (1981)	Californians having health checkups, 176 female twins and 91 males	No	8 ppm CO	0.6 (1/181)	
Jarvis et al. (1987)	Clinic outpatients	No	10 ppm CO (expired air)	16 (19/121)	
		No	1.7% CO (Hb)	18 (22/121)	
Bauman, Koch, Bryan (1982)	Adolescent nonclinic setting	No	6 ppm CO (expired air)	0	
			8 ppm CO (expired air)	3	
Stookey et al. (1987)	Cessation study	Yes	8 ppm	0 (0/20)	
Fortmann et al. (1984)	Representative sample for cardiovascular risk study	No	8 ppm	4.2 (37/890)	
Part II. Intervention samples					
Delarue (1973)	Canadians attending voluntary antismoking clinic	Yes	2% COHb	20.6 (22/107)	1-yr followup
			4% COHb	9.3 (10/107)	
			6% COHb	4.7 (5/107)	

TABLE 2—Continued

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Ohlin, Lundh, Westling (1976)[a]	Swedish patients with smoking-related diseases attending antismoking clinic and given nicotine gum	Yes	0.8% COHb	19.2 (25/130) 32.1 (35/109)	19.2% false reports at 1-wk followup; 32.1% false reports of not smoking at 6-mo followup
Isaesson and Janzon (1976)	Swedish heavy smokers in quit-smoking research project	Yes	1% COHb	8.8 (3/34)	8–9 wk followup
Lando (1982)	US smokers in multigroup smoking cessation study	Yes	CO	0 (0/22 to 60)	1 yr followup
Malcolm et al. (1980)[a, b]	UK trial of nicotine chewing gum	Yes	1.6% COHb	41.6 (47/113)	1-mo followup
Raw et al. (1980)	UK smokers attending a smokers' clinic in comparison of psychologic treatment and use of nicotine gum	Yes	CO or COHb	0 (0/33)	1-yr followup
Lando (1981)	US smokers in multigroup smoking cessation study	Yes	CO	Between 1.4 (1/74) and 4.2 (1/24)	Not clear when 1 "deceiver" withdrew from study; 1-wk (1/74) to 1-yr (1/24) followup; abstinence status also based on reports of informants
Jarvis et al. (1982)	UK smokers attending a smokers' clinic in trial of nicotine gum	Yes	CO or COHb	0 (0/26)	1-yr followup

TABLE 2.—Continued

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Russell et al. (1987)	UK smokers attending general practitioners	Some groups	7 ppm CO	About 22	4-mo to 1-yr followup
Glasgow et al. (1984)	US worksite smoking control study	Yes	10 ppm CO	0 (0/4)	6-mo followup
Jamrozik, Fowler et al. (1984)[a]	UK smokers in trial of nicotine gum	Yes	12 ppm CO	28.0 (7/25)	6-mo followup
Clavel et al. (1985)	French trial of acupuncture and nicotine gum	Yes	5 ppm CO	0 (0/24)	Sample of study participants (N=24); 1-yr followup
Lando and McGovern (1985)	US subjects undergoing various treatments for eliminating smoking	Yes	CO	2 cases out of at most 90	Up to 2-mo followup
Richmond and Webster (1985)	Australian smokers in a general practice; randomized trial of effects of advice to give up	Test group	COHb, SCN⁻, cotinine in plasma, and reports by family and friends	5.7 (2/35)	Criteria not stated; 6-mo followup
Abrams et al. (1987)	Worksite cessation	Yes	<9 ppm CO (expired air)	11.1 (1/9)	8-wk followup
Glynn, Gruder, Jegerski (1986)	Chicago Lung Association cessation study	Yes	10 ppm CO (expired air)	15.6 (7/45)	4-wk followup
Part III. High-risk/medical patients					
Li et al. (1984)	US asbestos-exposed smokers receiving (1) behavioral counseling or (2) minimal warning	Yes	9 ppm CO	1: 22.2 (4/18) 2: 23.1 (3/13)	11-mo followup

TABLE 2.—Continued

Reference	Population	Told to give up	Criterion for false reports of not smoking	% (n/N) False reports	Comments
Vogt et al. (1977)	San Francisco Center of MRFIT	Yes	8 ppm CO	4.4 (2/45)	
Sillett et al. (1978)[a]	UK study in 2 groups: (A) survivors of MI and (B) volunteers in nicotine gum trial	Yes	1.7% COHb	A: 21.6 (11/51) B: 40.2 (33/82)	
Ronan et al. (1981)	Irish post-MI patients	Yes	1.6% COHb	8.8 (5/57)	Mean 8.6-yr followup
Research Committee of the British Thoracic Society (1983)[a]	UK patients with smoking-related diseases in 4 group intervention trials involving advice, booklet, placebo, and nicotine polacrilex gum	All groups	1.6% COHb and 73 µmol/L SCN in plasma	27 25	27% false reports rate at 6-mo followup; 25% false reports rate at 1-yr followup

NOTE: CO=carbon monoxide; n/N=number of individuals reporting not smoking but with levels of biochemical marker exceeding cutpoint divided by all individuals reporting not smoking; COHb=carboxyhemoglobin; ppm=parts/million; SCN^-=thiocyanate; MRFIT=Multiple Risk Factor Intervention Trial; MI=myocardial infarction.

[a]Studies classified as outliers due to low criterion cutoffs.

[b]May be same group as (B) in Sillett et al. (1978).

SOURCE: Adapted from Lee (1988).

percent for either cotinine or CO. For intervention studies, values are typically 2 to 5 percent for cotinine and 0 to 10 percent for CO. High risk/medical samples appear to have the highest rates of misclassification of former smokers with the rates exceeding 20 percent. For example, as shown in Table 1, Jarvis and colleagues (1987) reported very low rates (1 percent) of false reporting in vascular patients who were not advised to quit compared with the rate in high-risk patients who were advised to quit (17 percent). It is likely that the pressure to stop smoking influenced the accuracy of patient reporting.

Observation studies in which no intervention occurs, or intervention studies in which there is minimal intervention or interaction with smokers, are less likely to prompt false reports of smoking cessation than studies in which intensive intervention does occur. In the former types of studies, in which no or low-intensity intervention occurred, there was a much lower prevalence of subjects reporting a 24-hour quit attempt during the past 6 months or current abstinence (Prochaska et al. 1985) than in intensive intervention studies, making misreporting less likely. A greater tendency to misreport in no or low-intensity intervention studies might occur with adolescents, for whom pressures to report not smoking may be omnipresent (Pechacek, Murray et al. 1984; Chapter 2, see section on Bogus Pipeline). A similar pressure might occur in some other instances, such as worksites in which a ban has been placed on smoking, where no intervention occurs but there may still be pressure on individuals to misreport. However, no studies have looked at the possibility of misreporting in such instances. The context in which the study takes place is likely to influence the degree of misreporting. Data currently being collected from smoking cessation programs in a wide variety of contexts may help to clarify this issue.

Clinic interventions and intensive interventions, on the other hand, typically ask participants to set a quit date. Close relationships are developed with the counselors, and self-reports of quitting are often given initially in a peer group. Under these higher demand conditions, biochemical verification may be needed to decrease the misreporting of current smokers as former smokers. For example, in MRFIT, special intervention subjects claiming to be former smokers at followup examinations had mean SCN^- levels between those of never smokers and continuing smokers (Ockene et al. 1982). Similar discrepancies between reported and validated cessation rates did not occur for the usual care men who had not received intensive intervention.

The use of biochemical tests for validating self-reports in epidemiologic studies has a number of limitations. The tests do not have perfect sensitivity and specificity; their half-lives do not necessarily fit the timeframe to be covered; and not all subjects are willing to provide the necessary samples for assessment. A very sensitive test may misclassify subjects as smokers if they have heavy passive smoke exposure (DiGuisto and Eckhard 1986; Haddow, Palomaki, Knight 1986; Haley et al. 1989; Jarvis et al. 1985), smoke occasionally (i.e., 1 or 2 cigarettes on isolated occasions) (Williams et al. 1979), and/or use nicotine in some other form, such as nicotine polacrilex gum or smokeless tobacco (Cohen et al. 1988; Slattery et al. 1989). Biochemical markers are also limited because they assess relatively short-term cessation (less than 2 weeks), and in studies concerned with the impact of cessation on health, there is more interest in evaluating consequences of long-term cessation.

In large-scale studies, use of biochemical assessments is generally not feasible; thus, mandatory use of such assessments and subsequent classification of refusers as smokers (as suggested by some investigators involved in clinical intervention studies e.g., Windsor and Orleans 1986) would result in an unacceptable distortion of the outcome data. In addition, some subjects may drop out if validation is required. The effect of lost subjects on study results may be difficult to estimate. In contexts other than intensive intervention trials, self-reported smoking status at the time of measurement and concurrent biochemical assessment have been demonstrated to be highly concordant (Fortmann et al. 1984; Petitti, Friedman, Kahn 1981) (Tables 1 and 2). This high concordance supports the use of self-report as a valid measure of smoking status in observation studies of the health effects of smoking cessation.

PART II. ASSESSING THE CONSEQUENCES OF SMOKING CESSATION

Study Designs Used to Assess the Consequences of Cessation

Overview of Study Design

Most evidence on the health benefits of smoking cessation derives from studies of human populations and not from animal studies or other types of research. Research on humans can be classified as experimental (the investigator assigns subjects to be exposed or not exposed to the risk factors or preventive factors of interest) or observational (the investigator does not determine whether subjects are exposed or not exposed to the factors of interest; exposure reflects the subjects' choices or some other process). Intervention studies include randomized or nonrandomized community-based investigations and clinical trials. The clinical trial, involving randomization of subjects to be exposed or not exposed to an intervention, has been used to investigate the effects of smoking cessation in patient groups and in populations. The observational designs include the ecologic study, the cross-sectional study, the cohort study, and the case–control study.

The biases potentially affecting these studies can be broadly classified as selection bias, information bias, and confounding bias (Table 3) (Kleinbaum, Kupper, Morgenstern 1982). Selection bias refers to distortion of an exposure–disease relationship by the mechanism through which subjects are selected. Information bias arises from the incorrect categorization of subjects as exposed or not exposed or as diseased or not diseased. The resulting misclassification of subjects on exposure or disease status may occur in a random or nonrandom fashion (Chapter 2, Part I). Confounding bias refers to the distortion of the apparent effect of an exposure on risk caused by association with other factors that affect outcome (Last 1988). In the subsequent review of the study designs used to assess the benefits of smoking cessation, sources of bias most relevant to each design are highlighted.

TABLE 3.—Examples of potential methodologic problems in investigating the health consequences of smoking cessation

Problem	Consequences
Current smokers developing symptoms of disease quit smoking	Apparent benefits of cessation are reduced
Self-reported former smokers are actually smoking (information bias)	Apparent benefits of cessation are reduced
Former smokers tend to have smoked less than persistent smokers (confounding bias)	Failure to account for the difference may exaggerate the apparent benefits of cessation
Former smokers tend to have a healthier lifestyle than persistent smokers (confounding bias)	Failure to account for the difference may exaggerate the apparent benefits of cessation
Smoking practices and the presence of smoking-related diseases affect participation in studies (selection bias)	Apparent benefits of cessation may be increased or decreased
Small number of subjects in a study	A beneficial effect of cessation may not reach statistical significance

Ecologic Studies

Ecologic studies represent a descriptive approach for examining the relation between risk factors and disease. Groups, rather than individuals, are the unit of analysis in ecologic studies. For example, changes in lung cancer mortality rates for selected countries have been examined for correlation with changes in measures of smoking for those countries, such as the percentage of smokers or per capita cigarette consumption (US PHS 1964; Cairns 1975; Cummings 1984; Doll and Peto 1981). Ecologic studies often have the advantage of being performed inexpensively and feasibly by using already available data. This design has well-described limitations related to the estimation of exposure and control of confounding, and may yield seriously biased data on exposure–disease relationships (Kleinbaum, Kupper, Morgenstern 1982; Rothman 1986).

Cross-Sectional Studies

In a cross-sectional or prevalence study, exposure and outcome are assessed at the same point in time among individuals in a population. Because cross-sectional studies measure exposure and outcome variables simultaneously, the true temporal relation between exposure and disease may be obscured (Rothman 1986). However, cross-sectional studies can be readily performed and have supplied much of the evidence on smoking cessation and nonmalignant respiratory diseases (Chapter 7).

Cross-sectional studies may be affected by selection bias. Because cigarette smoking is a strong cause of disease and death, groups studied cross-sectionally may not accurately reflect the natural history of smoking, smoking cessation, and the development of smoking-related illness. The proportion of heavier smokers and more susceptible smokers may be reduced compared with the original birth cohorts giving rise to the cross-sectional study population (McLaughlin et al. 1987). Former smokers who stopped because of the development of disease may be underrepresented, whereas those who stopped to reduce the risk of illness may be overrepresented.

Information bias is also of potential importance in cross-sectional studies. Preexisting conditions in survey participants may affect recall of past smoking or may alter the approach used by interviewers to gather smoking information. However, as summarized in Tables 1 and 2, cross-sectional surveys generally demonstrate low rates of misreporting of smoking status when compared with cotinine and CO levels.

As mentioned previously, a single observation on smoking behavior may lead to misclassification of smokers because of the dynamic nature of smoking behavior. Former smokers are typically a heterogeneous group with periods of abstinence ranging from days to years. For example, in the 1986 Adult Use of Tobacco Survey (US DHHS 1989), the subjects' responses were classified in 10 categories, 4 of which included former smokers. Of the former smokers, 12.5 percent had quit within the past 3 months, 7.8 percent had quit in the past 3 to 12 months, 22.3 percent had quit in the past 1 to 5 years, and 57.4 percent had quit 5 or more years earlier.

Cohort Studies

In a cohort study, the subjects are selected on the basis of exposure status (e.g., smoking behavior) and observed for development of disease. Observation may be forward in time (prospective), backward in time (historical or retrospective), or both. Correct conclusions can usually be made about the temporal relation between exposure (smoking cessation) and outcome (reduction of morbidity or mortality). With the cohort design, multiple health outcomes can be considered simultaneously. For example, the CPS-I and CPS-II conducted by the American Cancer Society (ACS) examined the effect of smoking behavior on total mortality and specific causes of death.

In a study of smoking cessation, selection bias could affect the findings of cohort studies if subjects lost to observation were more or less likely to benefit from smoking cessation than subjects remaining under observation (Greenland 1977). For intervention studies and cohort studies, the rate of subject loss provides an index of the potential selection bias.

In a cohort study of smoking cessation, some misclassification of exposure may be introduced if the classification of smoking status is based on a single assessment. Although the categorization of smoking status may be correct at the time the information is collected, inevitably some former smokers will resume smoking and some current smokers will stop. The extent of the resulting error will increase with the duration of followup. The resulting misclassification will tend to underestimate the effects of quitting because those who relapse to become current smokers would not be expected to experience beneficial effects attributable to quitting.

For example, in ACS CPS-I involving nearly 1 million people, Hammond and Garfinkel (1969) studied changes in smoking status over a 2-year period. Male former cigarette smokers in 1959–60 who reported that they were smoking in 1961–62 varied according to duration of prolonged abstinence reported in the 1959–60 survey. For respondents abstinent less than 1 year in 1959–60, 37.3 percent reported smoking 2 years later; of those reporting abstinence for 1 to 2 years, 19.1 percent were smoking 2 years later; and of those reporting abstinence of more than 2 years, 4.6 percent were smoking 2 years later. For all males who were former smokers in 1959–60, 11.3 percent reported smoking 2 years later. For all female former smokers in 1959–60, 6 percent reported smoking 2 years later. In the U.S. Veterans Study (Rogot and Murray 1980; Kahn 1966), male veterans in a cohort of 248,846 were classified based on responses to questionnaires administered in 1954 or in 1957 (if the 1954 questionnaire was not returned) and then followed for 16 years to determine the relationship between tobacco use and mortality. Undoubtedly, many of the original current smokers became former smokers as a result of the strong trend of smoking cessation among U.S. males during the followup period (US DHHS 1989).

Repeated assessment of smoking status in a cohort study can mitigate misclassification due to changes in smoking status over time (Chapter 2, Part I). Repeated measures are often feasibly made in cohort studies to minimize the effects of misclassification. Alternatively, validation substudies can be conducted within the cohort to quantify misclassification errors (Greenland 1988).

Case–Control Studies

Case–control studies involve selection of study subjects based on the presence (cases) or absence (controls) of a disease. Exposure and other attributes of cases and controls (e.g., smoking status or lifetime cigarette consumption) are then measured. The groups are compared with respect to the proportion having the attribute of interest to calculate the exposure odds ratio, which estimates the relative risk associated with exposure. Case–control studies can generally be conducted in less time than cohort studies or intervention studies and are less expensive to perform. Case–control studies are well suited for evaluation of diseases with low incidence rates.

Case–control analyses may be affected by information bias and selection bias. Case–control studies are prone to information bias if lifetime exposure histories are collected by interview (Schlesselman 1982). Retrospective lifetime histories of smoking or other exposures obtained from ill or elderly subjects may introduce misclassification. Similarly, studies that rely on reports from surrogates to assess smoking may misclassify exposure. If individuals classified as cases recall more accurately or less accurately than those classified as controls, differential misclassification results (Gordis 1982). Differential misclassification may also be introduced if respondents deliberately falsify answers or if interviewers differentially gather information from cases and controls (interviewer bias); interviewers not blinded to case–control status may probe more intensely for a putative causal exposure in cases than in controls (Sackett 1979). Blinding is often not feasible, and meticulous attention must be directed to training interviewers and to designing questionnaires to remove the possibility of interviewer

bias. Although selection bias may affect any case–control study that is not population-based, it is unlikely to be of particular importance in most case–control studies of smoking cessation.

Intervention Trials

Intervention trials are designed to test a hypothesized cause–effect relationship or the benefits of a preventive program by modifying the putative causal or preventive factor and measuring the effect on relevant outcome measures. Intervention trials may be directed at individuals or groups, such as communities. Regardless of the unit of observation, the trials may be conducted with (e.g., a clinical trial) or without randomization to the intervention.

Clinical trials are most commonly used to assess therapeutic interventions, but this design has also been used to evaluate preventive interventions, such as smoking cessation. A clinical trial includes one or more comparison groups in which subjects receive the control intervention; subjects are randomly assigned to the treatment and comparison groups to ensure that the groups are comparable with respect to characteristics potentially affecting the outcomes of interest. Individuals or groups such as communities can be the units of randomization. Within the limits of chance, random assignment makes the intervention and control groups similar at the onset of study.

Although widely used to test smoking cessation methods, clinical trials have been used infrequently to assess the health benefits of smoking cessation. In comparison with observation studies, the clinical trial design offers the potential for eliminating or more tightly controlling bias from the selection of subjects and from confounding. However, for many health outcomes, both a large sample size and a lengthy followup period may be needed to have sufficient statistical power. Moreover, in a study of smoking cessation, the power of the trial also depends on the extent of the reduction in smoking in the intervention group, in comparison with the control group. In the reported smoking intervention trials, only a minority of participants attained continuous or prolonged abstinence following most cessation interventions (Hunt, Barnett, Branch 1971; Hunt and Bespalec 1973; Ockene et al. 1990). Even with intensive, prolonged interventions, as in MRFIT, only 42 percent of smokers within the special intervention group were not smoking at 6-year followup, and only 26 percent of baseline smokers had been continuously abstinent from cigarettes over this prolonged period (Ockene et al. 1990).

Only a few clinical trials provide information relevant to the health benefits of cessation (Chapter 3). In the Whitehall Civil Servants Study (Rose et al. 1982), the investigators randomly intervened in smoking with advice from a physician in a group of men at high risk for cardiopulmonary disease. In MRFIT, smoking intervention was one component of the risk factor intervention program directed at the special intervention group (MRFIT Research Group 1982).

In most clinical trials that assess the effect of cessation on disease outcomes, such as the Whitehall Civil Servants Study (Rose et al. 1982), the investigators did not monitor longitudinally the persistence of quitting or levels of biochemical markers. The only clinical trial that has provided these measures is MRFIT (Ockene et al. 1990). Although