# Approaches for Conducting Large Cohort Studies

Walter C. Willett and Graham A. Colditz

## INTRODUCTION

Numerous cohort studies have been initiated over the past 50 years to address lifestyle and biochemical determinants of specific diseases. A key feature of these studies is that information on the primary exposures and potential confounders is required from the individual subjects themselves, rather than from routinely collected data. In this brief overview we will discuss some of the theoretical and practical issues to be considered in the development and conduct of a large modern cohort study.

Earlier cohorts, such as those of occupational and medical exposures including the health effects of ionizing radiation (1, 2), frequently did not require contact with the participants, and, thus, could address only very specific issues and were limited in ability to account for confounding. These studies will not be considered in detail here. Some of the first modern cohort studies were established specifically to evaluate one or a few factors, such as the health effects of smoking (3–5) or physical activity (6, 7). Still, other simple variables were collected that have provided valuable opportunities to examine the effects of obesity (8), alcohol consumption (9, 10), and oral contraceptives (11, 12). Early population-based cohorts included the work of Archie Cochrane who, in 1950, began the study of Rhondda Fach which encompassed eight mining towns. Of the adult population, 89 percent were x-rayed, including some 2,764 miners and 3,055 ex-miners (13, 14). Though the initial focus was on causes of pneumoconiosis and progressive massive fibrosis among miners living in these communities (14), rheumatoid arthritis was examined (15), and by the first follow-up 2½ years after the baseline data collection, psychiatric conditions were added as a focus of disability (16). Cardiovascular risk factors and endpoints were studied from early in the life of this cohort. Methodological work addressed the measurement and classification of coronary disease for epidemiologic studies (17). Miall and colleagues (18) described distributions of blood pressure, and identified risk factors for hypertension.

In the 1950s, the Washington County cohort was established by Comstock and colleagues (19) following a census of the county, and the Tecumseh Community Health Study was begun in 1957 to follow a whole community relating lifestyle to health and disease (20, 21). Conditions under study included coronary heart disease, hypertension, chronic obstructive lung disease, diabetes mellitus, arthritis, and obesity. Familial aggregation of disease was a focus from the inception of the study. Other population-based studies begun in the 1950s included the Framingham Heart Study (22), the Alameda County (California) study of health (23), and the study of cardiovascular disease in three southern US communities, Evans County, Georgia, Charleston County, South Carolina, and North Carolina (24).

Studies of the natural history of lung function and pulmonary symptoms required that the population would need to be readily accessible and stable to facilitate follow-up examinations and high participation. Fletcher established a cohort of 1,163 men in 1961 who were followed at 6 monthly intervals until 1969 (25). Additional populations were established for study of lung function in Berlin, New Hampshire, and East Boston, Massachusetts.

More recently, large cohort studies have been developed that attempt to attain a more comprehensive assessment of lifestyle, in addition to medical and familial factors, to address all potential determinants of specific diseases in a comprehensive manner. Such cohorts usually do not aim primarily to maximize the contrast in exposure but, rather, draw on general populations. An exception to this has been the recently created National Cancer Institute cohort to study diet and cancer in older Americans in which the population was screened for fat intake and the extremes were over-sampled. Because of the likely importance in the etiology of many diseases, these general purpose cohort studies have included a comprehensive assess-

ment of diet. At least 30 such large prospective studies that include assessment of diet have now been established (see table 1). These are mainly in Europe and North America, but several cohorts have now been initiated in Asia. Together, these represent substantial ethnic diversity and include approximately 3 million persons under observation.

Because the establishment of a cohort typically represents a major investment in resources, it is efficient to examine all major health outcomes, including cancers and cardiovascular diseases. Common less lethal conditions, such as gallstones or kidney stones, can also be examined within cohorts, even though it would often be impossible to justify funding to develop a cohort de novo to examine these endpoints. More recently, issues of quality of life and cognitive function are being addressed in such studies. Because funding organizations are typically structured for a specific disease or organ system, a broad view of the ultimate use of the cohort is critical at its inception. The ability to evaluate multiple endpoints simulta-

neously within the same population is particularly valuable when there are competing risks and benefits. For example, moderate consumption of alcohol has been found to increase risk of breast cancer and reduce risk of coronary heart disease, raising the question about the overall impact on health. In this situation, an examination of total mortality or a composite endpoint of death or all major illness can be particularly useful, though the risk/benefit trade-off may still depend on the underlying distribution of major risk factors such as cigarette smoking. Also, from a methodological standpoint, the documentation of a relation between an exposure and one or more outcomes makes the lack of association with another outcome within the same cohort more compelling. For example, the documentation of inverse associations between dietary fiber and risks of diverticular disease and coronary heart disease within the Health Professionals Follow-up Study (26), made it unlikely that the lack of clear association with colon cancer (27) was due to inadequate variation in exposure or imprecision in measuring fiber intake.

**TABLE 1.    Large prospective studies of diet and disease using comprehensive food frequency questionnaires***

| Study | Country | Population | No. | Started | Biologic specimens |
|---|---|---|---|---|---|
| Israeli Ischemic Heart Disease Study | Israel | Males | 10,000 | 1963 | None |
| Norwegian Cohort Study | Norway | Males and females | 17,000 | 1967 | None |
| Adventist Health Study | United States | Males and females | 40,000 | 1976 | None |
| Nurses' Health Study | United States | Females | 90,000 | 1980 | 33,000 bloods, 68,000 nails |
| Canadian Breast Screening Study | Canada | Females | 57,000 | 1982 | None |
| New York University Women's Health Study | United States | Females | 14,000 | 1985 | 14,000 bloods, repeated samples |
| The Alpha-tocopherol Beta-carotene Cancer Prevention Study | Finland | Males | 29,000 | 1985 | 29,000 bloods |
| Northern Sweden Health and Diseases Study | Sweden | Males and females | 55,000 | 1985 | 50,000 bloods |
| Health Professionals Follow-up Study | United States | Males | 52,000 | 1986 | 18,000 bloods |
| Iowa Women's Health Study | United States | Females | 42,000 | 1986 | None |
| Netherlands Cohort Study | Holland | Males and females | 121,000 | 1986 | None |
| Sweden Mammography Cohort | Sweden | Females | 61,000 | 1987 | None |
| Atherosclerosis Risk in Communities | United States | Males and females | 16,000 | 1987 | 16,000 bloods |
| ORDET† | Italy | Females | 11,000 | 1987 | 11,000 bloods, urine, and nails |
| Honolulu Heart Program | United States | Males (Japanese) | 8,000 | 1988 | Blood |
| JPHC† | Japan | Males and females | 140,000 | 1990 | 49,000 bloods |
| Melbourne Collaborative Cohort Study | Australia | Australian-born Italian and Greek migrant males and females | 42,000 | 1990 | 42,000 bloods |
| Nurses' Health Study II | United States | Young female nurses | 95,000 | 1991 | ~35,000 bloods plus urines‡ |
| American Cancer Society | United States | Males and females | 184,000 | 1992 | None |
| Canadian Study of Diet, Lifestyle, and Health | Canada | Males and females | 100,000‡ | 1992 | Hair plus toenails for 97% |
| Women's Health Study§ | United States | Females | 40,000 | 1992 | 27,000 bloods |
| EPIC† | 9 European countries | Males and females | 440,000 | 1993 | 350,000 bloods |
| National Cancer Institute | United States | Males and females | 540,000 | 1995 | None |
| Multi-Ethnic Cohort | United States | Males and females (multiethnic) | 215,000 | 1993 | Blood and urine on subsets |
| Singapore Cohort Study | Singapore | Males and females (Chinese) | 48,000‡ | 1993 | Blood and urine on subsets |
| Women's Health Initiative§ | United States | Females | 165,000‡ | 1993 | 164,500 bloods, ~10,000 urines |
| Women's Antioxidant Cardiovascular Study§ | United States | Females | 8,000 | 1994 | 5,800 bloods |
| California Teacher's Study | United States | Females | 132,000 | 1995 | None |
| Black Women's Health Study | United States | Black females | 65,000 | 1995 | None |
| Growing up in the '90s | United States | Male and female adolescents | 15,000 | 1996 | None |

* Modified from Willett (36).
†ORDET, Prospective Study of Hormones and Diet in the Etiology of Breast Cancer; JPHC, Japan Public Health Center-based Prospective Study; EPIC, European Prospective Investigation into Cancer and Nutrition.
‡ Planned number, still enrolling.
§ Dietary data collected within a randomized trial.

Because modern cohort studies provide the capacity to measure a wide variety of both exposures and outcomes, they provide the opportunity to address a large number of important public health issues. To make maximal use of such a resource requires the involvement of numerous investigators, both because of the volume of potential issues and because specialized knowledge in various substantive areas is required. Thus, the modern cohort study has been the analog of the massive particle accelerator in nuclear physics, whereby only a limited number of such resources can be supported, but many scientists cooperate to maximize the use of the facility. This is rapidly changing the nature of epidemiologic research, from numerous small fiefdoms to cooperative groups of interacting investigators. Such working relations, although created by necessity, can greatly enhance the level of scientific endeavor and productivity.

In this review we discuss some details of study design, including selection of the population, methods of data collection, follow-up and mortality surveillance, the benefits of updating exposure information, common methods for disease ascertainment and documentation, analytical issues that are pertinent, and the benefits of long term follow-up for refinement of our understanding of disease pathogenesis. We draw heavily on experience with the Harvard cohort studies because we are most familiar with their operations, but emphasize that many of the methods are applied to other settings.

## SELECTION OF POPULATION

A critical factor to consider in the development of a large cohort study is the trade-off of greater variation in exposure that may be present in a general population against the reduced potential for confounding and higher data quality that may result from selecting from a restricted group. For example, Doll and Hill recruited British physicians to study the relation between smoking and lung cancer (28), which ensured high quality information on smoking and other lifestyle factors. This also reduced the variation in socioeconomic status, education, occupation, and other factors that might have theoretically confounded the smoking-cancer relation.

Feasibility of follow-up is a key consideration, as high follow-up is essential for the internal validity of the study results. The British Doctors Study was designed to draw on the medical registration board which records all deaths among British doctors, hence facilitating a follow-up scheme that is low cost and highly accurate. Once deaths were identified, additional details were required to confirm the cause of death (see below). Another advantage of professional groups is

the presence of listings of potential participants. Such listings facilitated the creation of the British Doctors Study, the Nurses' Health Study, the Health Professionals Follow-up Study, the California teachers cohort, and the British civil servants cohort. Likewise, the Dorn study of some 300,000 US veterans who were active holders of government life insurance policies, creatively used insurance claims to the Veterans Administration as a means to ascertain mortality in the cohort (29). The Iowa Women's Study, on the other hand, drew on the availability of a computerized listing of women with a driver's license to initiate the cohort study of diet and cancer in women 50–65 years of age living in Iowa. One advantage of the use of professional groups as the source of cohort populations has been the high level of accuracy for data recorded by participants. An added advantage of cohorts based on health professionals has been the ease of obtaining biologic specimens at low cost.

The baseline recruitment strategy can be tailored to increase the rate of subsequent follow-up. However, hurdles to initial participation, such as the use of a longer questionnaire, do not typically impact the internal validity and may increase the likelihood of subsequent participation. In addition, it is important to include on the baseline questionnaire those key identifiers that may be required for subsequent searching should the participant become lost to follow-up. These items usually include the social security number and the name of a contact person who may be reached in the case of the investigators inability to locate a study participant. Other professional listings may help in this phase of a study. In the British Doctors Study, the use of the medical registration system to track registered physicians was an efficient method to trace cohort members. Registration listings of nurses and dentists have helped trace participants in more recent US cohort studies. With increasing computer processing power, it is possible to search state listings of registered nurses to locate study participants in the Nurses' Health Study who have moved and not provided their forwarding address to the investigators. This, of course, is more effective for those cohort members with rare family names.

## METHODS OF DATA COLLECTION

Most of the existing large cohorts have used a pencil and paper questionnaire format to collect baseline data. These data are processed and a computer database created. Advances in computer technology have facilitated the refinement of methods used in the ongoing cohort studies, though several were begun before high-speed computing allowed for easier data management and preceded technology such as national

telephone directories on CD-ROM that facilitate tracking of study participants who move. The early cohorts, such as the British Doctors Study (3), serve as the basis from which many of the refinements described in this review were developed.

Modern optical scanning technology has replaced double-key entry as the standard for high quality data entry. Estimates from the Nurses' Health Study using early optical scanning technology suggest that optical scanning may have one data error per 10,000 data points compared with three errors per 1,000 data points for double-key entry. In addition, optical scanning processes are now available at substantially lower cost than key entry. Though enhancing data quality, this approach puts a premium on multiple choice responses. Hence, pilot studies may be indicated with open-ended responses to determine the optimal categories. Because of the importance of including relevant covariates for all endpoints that might be examined in a general purpose cohort, it is useful to circulate the baseline questionnaire for review by peers at the onset of the study. This may help avoid omission of one or two important risk factors that, if omitted, may hinder analysis of a major endpoint.

Biologic samples have also been collected in several of the cohort studies underway. Clinic-based cohorts, such as the Framingham Heart Study (22, 30), the Normative Aging Study (31), the Tecumseh Community Health Study (21), and the Honolulu Heart Program (32, 33), added the collection of samples to their protocol as part of a routine follow-up visit. Because laboratory analyses are usually expensive on a per unit basis, large studies often draw on the cohort to conduct nested case-control studies with these biologic samples. This reduces costs and maintains study power (34).

Some large cohort studies begun in the 1980s did not computerize all baseline data but, rather, use the case-cohort approach to data analysis (35). This method reduces the cost of initial data processing.

## METHODS OF FOLLOW-UP

Once participants are recruited into the study, follow-up for identification of participants who develop major illnesses has used different methods. Alternatives have included follow-up questionnaires, linkage to tumor registries, and mortality surveillance. For example, in the Nurses' Health Study, follow-up questionnaires are mailed to all cohort members every 2 years. These questionnaires are sent with a cover letter and a newsletter that updates participants on the progress of the study. Each follow-up questionnaire inquires about a number of exposures as well as the development of cancer, cardiovascular disease, and other major medical conditions diagnosed since the

last follow-up. The first follow-up questionnaire is mailed in June of even-numbered years (1978, 1980, 1982, . . .), and those who do not respond are sent a second mailing in September. On average, 70 percent of the cohort respond to the first mailing. Subsequently, we send a second, third, and fourth questionnaire to those who still have not responded. Finally, a fifth mailing of a short questionnaire that includes only a few key exposure variables and the list of major illnesses is sent. This fifth mailing, which includes a newsletter to update participants, is sent in June of odd-numbered years. This ensures that any change of address is obtained from the postal service (the usual practice being to keep address forwarding orders for only 12 months). We have also used certified mail to raise the level of response to our short questionnaire in our cohort studies (36). In a randomized trial of five mailing strategies conducted within the Health Professionals Follow-up Study, Rimm and colleagues compared response rates to certified mail, United Parcel Service, window envelope with computer printed addresses, typewritten addressed envelope, and hand-addressed envelopes. Defining response rate as the return of a completed questionnaire, the highest response rate (63.2 percent) was observed for certified mail (36).

In 1982, we added a telephone follow-up to reach those women who had not responded to any of the five mailings. More than 14,000 women were successfully contacted and completed a brief telephone interview focused on a limited number of exposures and any newly diagnosed illness. Telephone follow-up was repeated after the 1986 follow-up cycle. The cost of phone follow-up is considerably higher than that of mailed questionnaires in the United States. In 1988, we used a series of additional approaches, including sending questionnaires by United Parcel Service and also by certified mail; these approaches achieved a response of 88 percent. In 1990 and subsequently, using both telephone and certified mail to reach initial nonresponders, responses to each follow-up cycle were received from just over 90 percent of the women in the study.

Each year we are notified of address changes by more than 3 percent of study participants. In addition, some mail is returned to us as undeliverable. Using mechanisms developed over the last 20 years, we trace these women through direct contact with the local postmaster, the State Boards of Nursing, and a contact person designated by the study participant (contacts were identified by study members in 1978, 1982, 1986, 1988, and again in 1992). We also mail to the occupant of the house last recorded as the residence for a lost participant, and may at times contact the

neighbors as an alternative source for a participant who has moved. Through these approaches, we have successfully located the majority of participants with whom we have lost contact at some time.

## UPDATING EXPOSURE INFORMATION

A powerful feature of cohort studies is the opportunity to collect repeated exposure data over time. In the Nurses' Health Study, for example, five comprehensive dietary assessments were collected between 1980 and 1994. Such repeat measurements of dietary intake provide many possible analytical opportunities to reduce the effects of random measurement error and to evaluate various hypothesized temporal relations between the dietary factor and the disease outcome (37). For example, some distinction between short- and long-term induction periods can be made using either the earliest or the most recent questionnaire, thus providing much greater power for short induction periods than would be available with only a baseline questionnaire. A more detailed evaluation of induction periods can be made by relating each repeated questionnaire to incidence of disease during specific categories of induction time (i.e., 1–2 years, 3–4 years, 5–9 years, . . .), and then combining the results for each specific category from all questionnaires. Measured changes in diets of individuals over time are a mix of true variation and measurement error. Thus, the comparison of persons whose intakes are consistently high with those whose intakes are consistently low can provide a strong test of cumulative exposure, as well as both long and short latency, because it is highly likely that these persons are truly high or truly low over long durations. An updated cumulative average of individual dietary intakes can make maximal use of all data for disease processes like atherosclerosis for which cumulative exposure is likely to be most relevant.

Because our understanding of disease etiology is often inadequate to specify a temporal relation with confidence, the use of several, rather than just one, analytical strategies to examine various temporal relations will generally be appropriate. If an association is observed, we can be more confident that a true relation exists if the relative risk becomes stronger with strategies that should provide maximum associations (such as when persons with consistently high intakes are compared with those having consistently low intakes). Also, clear evidence that an association is stronger with a particular temporal relation can provide important information on the pathogenic process and possibilities for intervention. For example, if the strongest relative risk is seen when only the initial measurement is used, and if the association gets stronger when early

follow-up is deleted, this would suggest that relatively short-term intervention trials might not be informative or that populations in older age groups might not be appropriate targets of prevention programs. If, on the other hand, no association is observed, the demonstration that this lack of relation is seen when a full range of temporal relations is examined provides the most compelling evidence that an important association has not been missed.

As examples of other lifestyle factors, estrogen use and cigarette smoking have been updated biennially in the Nurses' Health Study. This information was key to detecting the positive association between current use of estrogen and breast cancer, and noting that risk declines rapidly after stopping use (38). Likewise, for smoking, risk of coronary heart disease, and several cancers, is strongest among current smokers and declines rapidly with years since quitting (39–41). The constantly changing availability of postmenopausal hormone products and their patterns of use, such as the addition of progestins to postmenopausal estrogen and varying number of days per month that these products are used, would make a single assessment of replacement hormones at baseline rapidly irrelevant as a measure of current exposure. With more than 1 percent of smokers stopping every year, it is also necessary to update smoking status on a regular basis to estimate accurately the relation between both current smoking and disease as well as the benefits of quitting smoking.

Although repeated measures of exposure can provide great power and insight into biologic relations, they can also create practical and methodological challenges. Inevitably, some percentage of participants will not respond to each request for information, and after repeated cycles of follow-up, the percentage with less than complete data becomes substantial. Also, subjects who develop intermediate endpoints may alter their behavior due to their perceived evaluation in risk of the primary endpoint. For example, Shekelle et al. (42) found that persons diagnosed with high blood cholesterol changed their diet sufficiently to obscure a relation between dietary lipids and serum cholesterol. Methods to account for biases introduced by knowledge of immediate risk factors that are both theoretically sound and practical remain to be established. In some cases, ad hoc but reasonable approaches can be used. For example in analyses of hormone replacement therapy and breast cancer mortality, we stopped updating exposure information at the time of diagnosis of breast cancer, which is realistic because estrogen use is almost always discontinued at diagnosis, and used that date as the date of event for cancers that were eventually fatal (43).

Another benefit of repeated questionnaires is the ability to add items to the follow-up questionnaires to address new and evolving hypotheses. Among the many additions to the Nurses' Health Study have been some of the diseases and conditions examined in that study and variables such as diet, physical activity, waist and hip circumferences, voluntary weight loss, and screening behaviors. With active follow-up, it is usually possible to recruit study participants to contribute to substudies that estimate the level of measurement error associated with the questionnaire. Each major new exposure added to the Nurses' Health Study has been evaluated for these measures of performance (44, 45), as has been done in other settings. For example Kushi and colleagues evaluated the validity of self-reported measures of body composition in the Iowa Women's Study, and the assessment of voluntary weight loss within that cohort has been evaluated in detail (46, 47). In response to the emerging public health concern surrounding silicone breast implants, we were also able to add this exposure to assess the relation between implants and rheumatologic conditions, which had in large part been ascertained through earlier studies within the cohort (48). After confirming that the prevalence of implants was sufficiently high to justify the study, details of exposure were added to the 1992 follow-up.

## CONFIRMATION OF ILLNESSES

### Nonfatal illnesses

Among the methods used to confirm illnesses are linkage to cancer registries and review of medical record to confirm self-reported illnesses. In addition to simple confirmation, these records often provide detail on histology and other disease characteristics that cannot be obtained routinely from participants. While registries are helpful for cancer, they do not exist for most other major chronic diseases, hence this approach has limited applicability.

In the Harvard cohort studies, for any report of cancer (except basal cell skin cancer), we seek written permission from study participants to review their medical records. We telephone nonrespondents to obtain verbal confirmation of the information reported on the follow-up questionnaire (asking details of diagnosis and treatment such as chemotherapy). All medical records are reviewed by trained physicians, blinded to exposure information previously provided by the study participant.

For participants reporting a myocardial infarction or stroke, we also seek the medical records pertaining to the initial diagnosis. Myocardial infarction is classified as confirmed if the records meet the criteria of the World Health Organization, including symptoms and either typical electrocardiogram changes or elevations of serum cardiac enzymes (49). Stroke is classified according to the criteria developed by the National Survey of Stroke (50).

On the 1982 Nurses' Health Study questionnaire, we added an item seeking a history of fracture of the hip or forearm and details regarding the diagnosis of gallstones and cholecystectomy. Diagnostic details of these major medical conditions have been included on subsequent follow-up questionnaires. Using a similar approach, we have added documentation of self-reported colon polyps and a range of eye conditions, including cataract surgery, macular degeneration, and glaucoma.

After the 1984 follow-up questionnaire cycle, we mailed supplementary questionnaires to all women who had ever responded affirmatively to the question "Have you ever been diagnosed as having diabetes mellitus?" on any of the previous questionnaires. This supplementary questionnaire included items on symptoms of diabetes at the time of diagnosis, fasting and random glucose levels, oral glucose tolerance testing, presence of glycosuria or ketonuria, history of ketoacidosis (including hospitalization), history of diabetes treatment, and gestational diabetes.

Earlier cohort studies conducted in the United Kingdom to document the health consequences of cigarette smoking used populations of doctors (28) to reduce the likelihood of error in the reporting of illnesses and to facilitate follow-up. Similarly, in establishing a large cohort of women, a key consideration was the ability of participants to accurately report the diagnosis of major illnesses. Because most reported disease must be confirmed, even a small increase in false positive reports would substantially increase the cost of the study. The highly accurate reporting of major medical conditions by Nurses' Health Study participants has contributed greatly to the cost-effective nature of this large study.

After confirming illnesses reported on the 1978 and 1980 follow-up questionnaires, we assessed the level of agreement with medical records. Overall, almost all self-reported cancers were confirmed by medical record review (51). Application of strict criteria for cardiovascular endpoints may result in rejection of some true cases and a slight underestimate of the true incidence of disease, but with few false-positive diagnoses.

The reliability of reporting hypertension, high blood cholesterol, and fractures has been confirmed in random samples of women. Agreement between self-report and medical records has been high, more than 98 percent for these conditions. In contrast, for classic

connective tissue diseases, we were only able to document less than 20 percent of cases when applying standard diagnostic criteria as defined by the American College of Rheumatology to information contained in medical records.

For emerging research areas such as histologic classification of benign breast disease, we are able to obtain slides from women reporting a biopsy-confirmed diagnosis and use centralized review by expert breast histopathologists to apply uniform diagnostic criteria.

In sum, the need to attain high specificity in the diagnostic classification of cases under study is most important in cohort studies. Many, but not all, diagnoses thus will require review of medical records. Low sensitivity will not bias relative risk estimates as long as underascertainment is not related to exposure.

## Mortality follow-up

Mortality surveillance can include use of professional organizations when these track mortality (such as the British Doctors Study), the National Death Index (52), and the social security administration files, which have often been used for occupational cohort studies.

Most deaths in the Nurses' Health Study and Health Professionals Follow-up Study are reported by the subject's next of kin or by postal authorities, and are supplemented by searches of the National Death Index for deaths among the nonrespondents. Even though social security numbers were not initially obtained, using these methods we estimate that more than 98 percent of deaths in this cohort have been identified (52). We have obtained social security numbers in subsequent cohorts, which greatly facilitates use of the National Death Index.

## ANALYTICAL ISSUES

Because cohort studies are often ongoing, long-term undertakings, there is usually not a demarcated sequence of data collection, cleaning, and analysis. Instead, exposure measurements and incident endpoints continue to accrue, and deaths are continuously identified. Analyses need to be planned and conducted simultaneously with data collection, which can often create competition for study resources. As the number of endpoints that can be examined is large, decisions need to be made regarding priorities for analysis and the point in follow-up at which they should be analyzed. Many of these priorities will be determined by commitments and time lines related to funding. However, cohort studies tend to create analytical priorities in line with public health priorities because the most common causes of morbidity and mortality are those for which there will first be sufficient numbers of cases to analyze.

A more complicated issue is when to reevaluate a relation. For important or potentially important associations, the full picture almost never emerges with the first examination. The simple overall association is important, but if an association is seen, issues of dose and other aspects of exposure, temporal relations, and interactions often have more important practical implications. For example, in the case of postmenopausal hormone replacement therapy and breast cancer risk, no association with ever use was seen in a first examination within the Nurses' Health Study (53). However, with continued follow-up it was possible to examine current and past users with adequate power, and an association was seen with current use only (38). However, elevated risks were seen only for older postmenopausal women or those who used hormones for longer than 5 years, and little risk was seen with use for less than 5 years. During the follow-up it became common to prescribe progestins in combination with estrogen, and some believed this would reduce risk; however, combination therapy was found to be associated with at least the same elevation in risk as with estrogen alone (54). Other more detailed questions about the pattern of progestin use will require further follow-up. With extended follow-up it was possible to examine interactions in more detail, and it appears that the association between estrogen replacement is substantially stronger in women who have not gained weight after age 18 years, probably due to their lower baseline endogenous estrogen levels (55). Finally, with additional data, the relation between hormone replacement therapy and total mortality could be examined in an analysis that addressed the competing risks and benefits of these drugs (43). However, even after 18 years of follow-up, the balance of health effects associated with lifetime use of hormone replacement remains unclear. In this instance, in which a substantial number of new cases of breast cancer (now close to 1,000) are added every 2 years, repeated examination of this important issue at a short interval is appropriate. For less common conditions, it will often be appropriate to wait until the number of cases has increased by 50 to 100 percent.

## BIOMARKERS

Specimens have been collected in a number of ongoing cohort studies to further understand the mechanisms of disease etiology. The range of specimens has been broad and allows assessment of varying measures: plasma or serum (blood lipids, nutrients, hormones), red cells (adducts, some nutrients, and lipids), white cells and cheek cells (DNA analysis), nails (trace elements and heavy metal contaminants), and urine (hormones, some nutrients). With the addition of

these biomarkers, the most complete etiologic picture can be determined by assessing behavior (diet, activity, cigarette smoking, etc.), blood or tissue levels of the related biochemical factor, physiologic intermediaries (e.g., hormones, lipids), DNA (genetic effect modifiers), and characteristics of lesions (such as DNA mutations in tissue blocks from cancers and precursor lesions). In the area of nutritional epidemiology, however, the addition of biochemical indicators of diet is not likely to replace dietary questionnaires because practical indicators for integrated measures of most dietary intakes do not yet exist.

The rapidly evolving capabilities for analysis of archived DNA specimens create complex issues regarding informed consent and the need to convey results of analysis to participants. Because the testable hypotheses of interest are often not known at the time of DNA collection, only general consent for genetic analysis is possible without recontacting participants, which may sometimes be impossible or difficult. "Anonymization" by removing identification from data so even the investigators cannot determine the identity of specimens has been one approach to avoid breaches of confidentiality and the quandary of whether to notify subjects of their results. However, this can result in a loss of scientific value as linkage with other data becomes impossible. An important distinction has been between high penetrance genes, such as BRCA1 or BRCA2, and polymorphisms that are likely to convey only modest relative risks and have thus little direct clinical implication. In the latter case, a standard of practice has been that general consent is sufficient for analyses to proceed. At this moment, several national bodies are considering guidelines for use of DNA samples. For our own studies, we have established an external advisory committee to help guide our efforts in this area, in part because the implications of a particular test can change rapidly if specific treatment for a condition predicted by that test becomes available.

## LONG-TERM FOLLOW-UP

Long-term follow-up is required to allow a rigorous evaluation of the temporal relations between exposure and disease risk. Also, long-term follow-up may uncover important associations not observed in the short term. For example, in the follow-up of the atomic bomb survivors, the relation between age at exposure and risk of breast cancer has only become clear after many years (2). For most exposures, such analyses are greatly enhanced if exposure status is updated over time. In particular, data on temporal relations, together

with repeated measures that allow the assessment of change in exposure, is important for assessing the potential impact of preventive interventions.

## CONCLUSION

Longitudinal data available through cohort studies provide an important and often unique opportunity to define the relations between numerous aspects of lifestyle and health. The addition of biochemical and genetic measures has greatly expanded the potential for epidemiology to contribute to our understanding of disease processes and to identify potential interventions to prevent disease and disability. Continued refinement of exposure assessment and analytical methods will insure that observational epidemiology maintains a central role in public health.

## REFERENCES

1. McGregor H, Land CE, Choi K, et al. Breast cancer incidence among atomic bomb survivors, Hiroshima and Nagasaki, 1950–69. J Natl Cancer Inst 1977;59:799–811.
2. Tokunaga M, Land CE, Tokuoka S, et al. Incidence of female breast cancer among atomic bomb survivors, 1950–1985. Radiat Res 1994;138:209–23.
3. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits: a preliminary report. Br Med J 1954;1: 1451–5.
4. Dorn H. Tobacco consumption and mortality from cancer and other diseases. Public Health Rep 1959;74:581–93.
5. Hammond CE, Horn D. Smoking in relation to the death rates of one million men and women. In: Haenszel W, ed. Epidemiologic approaches to the study of cancer and other chronic diseases. National Cancer Institute monograph 19. Bethesda, MD: US Department of Health, Education, and Welfare, National Cancer Institute, 1966:124–204.
6. Morris JN, Kagan A, Pattison DC, et al. Incidence and prediction of ischaemic heart-disease in London busmen. Lancet 1966;2:553–9.
7. Paffenbarger RS Jr, Wing AL, Hyde RT. Physical activity as an index of heart attack risk in college alumni. Am J Epidemiol 1978;108:161–75
8. Lew EA, Garfinkel L. Variations in mortality by weight among 750,000 men and women. J Chronic Dis 1979;32: 563–76.
9. Kono S, Ikeda M, Tokudome S, et al. Alcohol and mortality: a cohort study of male Japanese physicians. Int J Epidemiol 1986;15:527–32.
10. Klatsky AL, Armstrong MA, Friedman GD. Alcohol and mortality. Ann Intern Med 1992;117:646–54.
11. Colditz GA. The Nurses' Health Study: a cohort of US women followed since 1976. J Am Med Womens Assoc 1995;50: 40–4, 63.
12. Colditz GA. Oral contraceptive use and mortality during 12 years of follow-up: the Nurses' Health Study. Ann Intern Med 1994;120:821–6.

13. Cochrane AL, Cox JG, Jarman TF. Pulmonary tuberculosis in the Rhondda Fach: an interim report of a survey of mining community. Br Med J 1952;2:843–53.
14. Cochrane AL, Miall WE, Clarke WG, et al. Factors influencing the radiological attack rate of progressive pulmonary fibrosis. Br Med J 1956;1:1193–9.
15. Miall WE, Caplan A, Cochrane AL, et al. An epidemiological study of rheumatoid arthritis associated with characteristic chest x-ray appearances in coal-workers. Br Med J 1953;2: 1231–6.
16. Barraclough B. In conversation with Kenneth Rawnsley: part I. Bull R Coll Psychiatrists 1988;12:2–12.
17. Higgins ITT, Cochrane AL, Thomas AJ. Epidemiological studies of coronary disease. Br J Prev Soc Med 1963;17: 153–65.
18. Miall WE, Oldham PD. Factors influencing arterial blood pressure in the general population. Clin Sci 1968;17:409–44.
19. Comstock GW, Abbey H, Lundin FE Jr. The nonofficial census as a basic tool for epidemiologic observations in Washington County, Maryland. In: Kessler II, Levin ML, eds. The community as an epidemiologic laboratory: a casebook of community studies. Baltimore, MD: Johns Hopkins Press, 1970:73–99.
20. Napier JA, Johnson BC, Epstein FH. The Tecumseh, Michigan Community Health Study. In: Kessler II, Levin ML, eds. The community as an epidemiologic laboratory: a casebook of community studies. Baltimore, MD: Johns Hopkins Press, 1970:25–46.
21. Higgins MW, Longini IM. Discussion: the Tecumseh Community Health Study. In: Rao DC, Elston RC, Kuller LH, et al., eds. Genetic epidemiology and coronary heart disease: past, present, and future. New York, NY: Alan R. Liss, 1984:43–5.
22. Dawber TR. The Framingham Study: the epidemiology of atherosclerotic disease. Cambridge, MA: Harvard University Press, 1980.
23. Hochstim JR. Health and ways of living: the Alameda County, California, population laboratory. In: Kessler II, Levin ML, eds. The community as an epidemiologic laboratory: a casebook of community studies. Baltimore, MD: Johns Hopkins Press, 1970:149–76.
24. Tyroler HA. Epidemiologic studies of cardiovascular disease in three communities of the southern United States. In: Kessler II, Levin ML, eds. The community as an epidemiologic laboratory: a casebook of community studies. Baltimore, MD: Johns Hopkins Press, 1970:100–22.
25. Fletcher C, Peto R, Tinker C, et al. The natural history of chronic bronchitis and emphysema: an eight year study of early chronic obstructive lung disease in working men in London. Oxford, England: Oxford University Press, 1976.
26. Aldoori WH, Giovannucci EL, Rimm EB, et al. A prospective study of diet and the risk of symptomatic diverticular disease in men. Am J Clin Nutr 1994;60:757–64.
27. Giovannucci E, Rimm EB, Stampfer MJ, et al. Intake of fat, meat, and fiber in relation to risk of colon cancer in men. Cancer Res 1994;54:2390–7.
28. Doll R, Hill AB. Lung cancer and other causes of death in relation to smoking: a second report on the mortality of British doctors. Br Med J 1956;2:1071–81.
29. Kahn HA. The Dorn study of smoking and mortality among US veterans on 8 years of observation. In: Haenszel W, ed. Epidemiologic approaches to the study of cancer and other chronic diseases. National Cancer Institute monograph 19. Bethesda, MD: US Department of Health, Education, and Welfare, National Cancer Institute, 1966:1–125.
30. Castelli WP, Abbott RD, McNamara PM. Summary estimates of cholesterol used to predict coronary heart disease. Circulation 1983;67:730–4.
31. Sparrow D, Borkan GA, Gerzof SG, et al. Relationship of fat distribution to glucose tolerance: results of computed tomography in male participants of the Normative Aging Study. Diabetes 1986;35:411–15.
32. Rhoads GG, Gulbrandsen CL, Kagan A. Serum lipoproteins and coronary heart disease in a population of Hawaii Japanese men. N Engl J Med 1976;294:293–8.
33. Kagan A, ed. The Honolulu Heart Program: an epidemiological study of coronary heart disease and stroke. Australia: Harwood Academic Press, 1996.
34. Muñoz A, Gange SJ. Methodological issues for biomarkers and intermediate outcomes in cohort studies. Epidemiol Rev 1998;20:29–42.
35. van den Brandt PA, van't Veer P, Goldbohm RA, et al. A prospective cohort study on dietary fat and the risk of postmenopausal breast cancer. Cancer Res 1993;53:75–82.
36. Rimm EB, Stampfer MJ, Colditz GA, et al. Effectiveness of various mailing strategies among nonrespondents in a progressive cohort study. Am J Epidemiol 1990;131:1068–71.
37. Willett WC. Nutritional epidemiology. 2nd ed. New York, NY: Oxford University Press, 1998.
38. Colditz GA, Stampfer MJ, Willett WC, et al. Prospective study of estrogen replacement therapy and risk of breast cancer in postmenopausal women. JAMA 1990;264:2648–53.
39. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation and time course of decreased risks of coronary heart disease in middle-aged women. Arch Intern Med 1994;154:169–75.
40. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation in relation to total mortality rates in women: a prospective cohort study. Ann Intern Med 1993;119:992–1000.
41. Fuchs CS, Stampfer MJ, Colditz GA, et al. Alcohol consumption and mortality among women. N Engl J Med 1995;332: 1245–50.
42. Shekelle RB, Stamler J, Paul O, et al. Dietary lipids and serum cholesterol level: change in diet confounds the cross-sectional association. Am J Epidemiol 1982;115:506–14.
43. Grodstein F, Stampfer MJ, Colditz GA, et al. Postmenopausal hormone therapy and mortality. N Engl J Med 1997;336: 1769–75.
44. Rimm EB, Stampfer MJ, Colditz GA, et al. Validity of self-reported waist and hip circumferences in men and women. Epidemiology 1990;1:466–73.
45. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol 1994;23:991–9.
46. Kushi LH, Kaye SA, Folsom AR. Accuracy and reliability of self-measurement of body girths. Am J Epidemiol 1988;128: 740–8.
47. French SA, Jeffery RW, Folsom AR, et al. Weight variability in a population-based sample of older women: reliability and intercorrelation of measures. Int J Obes Relat Metab Disord 1995;19:22–9.
48. Sanchez-Guerrero J, Colditz GA, Karlson EW, et al. Silicone breast implants and the risk of connective-tissue diseases and symptoms. N Engl J Med 1995;332:1666–70.
49. Rose GA, Blackburn H, Gillum RF, et al. Cardiovascular survey methods. 2nd ed. World Health Organization monograph series no. 56. Geneva, Switzerland: World Health Organization, 1982.
50. Walker AE, Robins M, Weinfeld FD. The National Survey of Stroke: clinical findings. Stroke 1981;12(Suppl 1):I13–44.
51. Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am J Epidemiol 1986;123:894–900.
52. Stampfer MJ, Willett WC, Speizer FE, et al. Test of the National Death Index. Am J Epidemiol 1984;119:837–9.
53. Buring JE, Hennekens CH, Lipnick RJ, et al. A prospective cohort study of postmenopausal hormone use and risk of breast cancer in US women. Am J Epidemiol 1987;125: 939–47.
54. Colditz GA, Hankinson SE, Hunter DJ, et al. The use of estrogens and progestins and the risk of breast cancer in postmenopausal women. N Engl J Med 1995;332:1589–93.
55. Huang Z, Hankinson SE, Colditz GA, et al. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997; 278:1407–11.