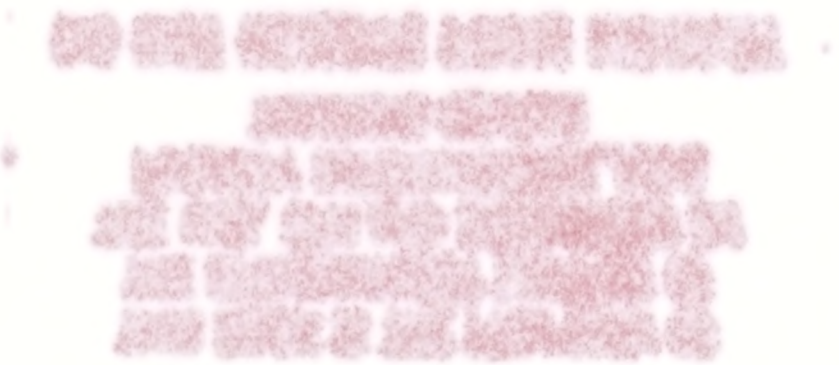
Nutritional Epidemiology

WALTER WILLETT



Nutritional Epidemiology

Monographs in Epidemiology and Biostatistics
edited by Brian MacMahon

1 THE EPIDEMIOLOGY OF DEMENTIA
James A. Mortimer and Leonard M. Schuman 1981

2 CASE-CONTROL STUDIES
Design, Conduct, Analysis
James J. Schlesselman 1982

3 EPIDEMIOLOGY OF MUSCULOSKELETAL DISORDERS
Jennifer L. Kelsey 1982

4 URBANIZATION AND CANCER MORTALITY
The United States Experience, 1950–1975
Michael R. Greenberg 1983

5 AN INTRODUCTION TO EPIDEMIOLOGIC METHODS
Harold A. Kahn 1983

6 THE LEUKEMIAS
Epidemiologic Aspects
Martha S. Linet 1984

7 SCREENING IN CHRONIC DISEASE
Alan S. Morrison 1985

8 CLINICAL TRIALS
Design, Conduct, and Analysis
Curtis L. Meinert 1986

9 VACCINATING AGAINST BRAIN DYSFUNCTION SYNDROME
The Campaign Against Rubella and Measles
Ernest M. Gruenberg 1986

10 METHODS IN OBSERVATIONAL EPIDEMIOLOGY
Jennifer L. Kelsey, W. Douglas Thompson, Alfred S. Evans 1986

11 CLINICAL EPIDEMIOLOGY: THE STUDY OF THE OUTCOME OF ILLNESS
Noel S. Weiss 1986

12 STATISTICAL METHODS IN EPIDEMIOLOGY
Harold A. Kahn and Christopher T. Sempos 1989

13 RESEARCH METHODS IN OCCUPATIONAL EPIDEMIOLOGY
Harvey Checkoway, Neil E. Pearce, Douglas J. Crawford-Brown 1989

14 CONCEPTION TO BIRTH
Epidemiology of Prenatal Development
Jennie Kline, Zena Stein, Mervyn Susser 1989

15 NUTRITIONAL EPIDEMIOLOGY
Walter Willett 1990

MONOGRAPHS IN EPIDEMIOLOGY AND BIOSTATISTICS VOLUME 15

Nutritional Epidemiology

Walter Willett, M.D.
Professor of Epidemiology and Nutrition
Harvard School of Public Health

New York Oxford
OXFORD UNIVERSITY PRESS
1990

Oxford University Press

Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Petaling Jaya Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland

and associated companies in
Berlin Ibadan

Published by Oxford University Press, Inc.,
200 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press

Library of Congress Cataloging-in-Publication Data
Willett, Walter.
Nutritional epidemiology / Walter Willett.
p. cm. — (Monographs in epidemiology and biostatistics ; v. 15)
Includes bibliographies and index.
ISBN 0-19-504501-7
1. Nutrition—Research—Methodology. 2. Nutrition surveys—
Methodology. 3. Diet in disease—Research—Methodology.
I. Title. II. Series.
[DNLM: 1. Diet. 2. Epidemiologic Methods. 3. Nutrition. W1
MO567LT v. 15 / QU 145 W713n]
QP141.W513 1989
614.5'939—dc20
DNLM/DLC for Library of Congress 89-9317 CIP

9 8 7 6 5 4

Printed in the United States of America
on acid-free paper

Preface

This book is written for those seeking to understand the relation between diet and long-term health and disease. A basic premise is that our understanding of biologic mechanisms remains far too incomplete to predict confidently the ultimate consequences of eating a particular food or nutrient. Thus, epidemiologic studies directly relating intake of dietary factors to risk of death or disease among humans play a critical complementary role to laboratory investigations. A great expansion of the literature in nutritional epidemiology and the development of a firmer quantitative basis for this science have occurred in the last ten years. Most importantly, the degree of variation in dietary factors among individuals has been quantified in many populations, standardized questionnaires have been developed for use in large epidemiologic studies, and the degree of accuracy of these questionnaires has been documented. Although most of the questions regarding the relations of dietary factors to disease remain unanswered, the foundations for obtaining this information have been laid. *Nutritional Epidemiology* represents an attempt to consolidate the substantial new methodologic information and experience that have been developed during the last decade.

I have written this book specifically for researchers actively engaged in studies of diet and disease and for persons seriously attempting to read and interpret published epidemiologic information relating to nutrition. The book is not intended for the casual reader seeking to learn what is known about the effects of diet and nutrition on human disease; such a work would rapidly be out of date. In addition, this book does not address issues related to child growth and studies of nutritional deficiency in developing countries; both are important topics but outside the scope of this text. However, I am confident that those involved in nutritional studies in developing countries can benefit from the material presented, as many of the principles are relevant.

Though I have not attempted to define or explain basic epidemiologic terms, most of the chapters can be read by someone with elementary statistical knowl-

edge but with no epidemiologic background. Readers without exposure to epidemiology would profit by referring to an introductory text such as MacMahon and Pugh's *Principles of Epidemiology*, Lilienfeld's *Foundations of Epidemiology*, Ahlbom and Norell's *Introduction to Modern Epidemiology*, Friedman's *Primer of Epidemiology*, Kahn and Sempos' *Statistical Methods in Epidemiology*, Hennekens and Buring's *Epidemiology in Medicine*, or Rothman's *Modern Epidemiology*—the latter being the most advanced text. Similarly, epidemiologists without formal exposure to nutrition can benefit by reading Guthrie's *Introductory Nutrition* and referring to standard texts such as Davidson and Passmore's *Human Nutrition and Dietetics* or Goodhart and Shils' *Modern Nutrition in Health and Disease*.

As the content of different chapters varies considerably in depth, many readers will want to reserve some sections for future reference rather than read the book from cover to cover. Chapter 1 provides an overview of nutritional epidemiology for those unfamiliar with the field; experienced epidemiologists may want to skim this material. Chapter 2 addresses the various perspectives from which diet can be viewed and discusses the calculation of nutrient intakes from data on food consumption. Chapter 3 provides a conceptual background regarding sources of variation in diet for the novice to this field and also assembles data on dietary variation that can be of use to the serious investigator. Chapter 4, written by my experienced colleague, Jelia Witschi, reviews the strengths, limitations, and appropriate applications of short-term recalls and dietary records as methods for measuring intake. Because of the central role of food frequency questionnaires in nutritional epidemiology, the design and evaluation of these questionnaires are examined in detail in Chapters 5 and 6. Recall of diet from the remote past is of potential importance for diseases with a long latency; the small but growing literature on this topic is reviewed in Chapter 7. The use of spouse surrogates for reporting of diet is common in case-control studies; this topic is covered in Chapter 8 by Jon Samet, who has contributed importantly to this issue.

Chapter 9, written by my former graduate student and present colleague, David Hunter, is divided into two parts. The first deals with conceptual and general issues in the use of biochemical indicators for assessing intake of specific dietary factors. The second part provides detailed information regarding the biochemical measurement of specific nutrients; many readers will probably want to use this material for reference. This chapter provides an epidemiologic perspective of the role of biochemical indicators in epidemiologic studies. Dr. Hunter points out that close collaboration between laboratory scientists and epidemiologists is essential when embarking on studies that involve biochemical measurement. Chapter 10 addresses the use of measures of body size and composition, also with a special emphasis on epidemiologic applications.

Chapters 11 and 12, in contrast with others, probe issues in the analysis and interpretation of epidemiologic data. Ironically, if the long-term effect of energy intake is being researched, measurement of total energy intake is usually irrelevant in an epidemiologic study—yet such a measurement is critical if nutrient composition rather than total energy intake is of interest. These issues are considered in detail in Chapter 11. The effect of measurement error in epidemio-

logic studies and statistical methods to compensate for measurement error are addressed in Chapter 12. This is an active area of development in epidemiology and statistics in general, but much of the stimulus has arisen from studies of diet. Although it is of general epidemiologic interest, I have included a chapter on this topic because of the important implications for nutritional studies and the lack of a practical overview for an epidemiologically oriented audience. This chapter, however, is intended for those with a substantial epidemiologic background.

Chapters 13 to 15 address substantive applications of nutritional epidemiology. They are intended to provide examples to reinforce the principles covered in the earlier chapters. The three topics selected, vitamin A and lung cancer, dietary fat and breast cancer, and diet and heart disease, are all of substantial current interest.

Many chapters on the relation of diet and disease remain to be written. The brief final section of this book attempts to anticipate the direction of future research and to encourage activity in areas that seem most promising.

Boston, Massachusetts W.W.
July 1989

To Gail

Acknowledgments

The concepts in this book owe much to the work and ideas of many predecessors, present colleagues, and former students. In particular, I am grateful for the encouragement, support, and ideas of my colleagues Frank Speizer, Meir Stampfer, Graham Colditz, Bernard Rosner, Laura Sampson, Frank Sacks, Simonetta Salvini, Mauricio Hernandez, David Hunter, Jelia Witschi, and Charles Hennekens at the Channing Laboratory and Harvard School of Public Health. Jon Samet provided strong initial encouragement to begin the book and provided further concrete support in the form of a chapter. Brian MacMahon first stimulated me to think seriously about diet as a possible cause of cancer, which led to much of the research underlying this book. Further development of ideas and approaches to the study of diet and disease have evolved from intellectual exchanges with Jim Marshall, Saxon Graham, John Potter, Dimitrios Trichopoulos, Peter Boyle, Richard Peto, and Geoffrey Howe.

Much of the material for this book was developed over the last ten years while I was teaching courses on nutritional epidemiology at the Harvard School of Public Health and the New England Epidemiology Institute. My insight and understanding as well as the teaching material have profited immensely from the ideas and comments of many students in these courses and from my doctoral students at Harvard School of Public Health.

The developmental work and data described in this book would not have been possible without a career development award and the funding of research grants, particularly the Nurses Health Study, from the National Institutes of Health. I am thus most grateful for the constructive comments of the many anonymous peer reviewers who have supported this work.

Susan Newman and Stephanie Bechtel provided critical support in the production of this book. Meir Stampfer, Graham Colditz, John Potter, and Matt Longnecker each reviewed the entire manuscript at one stage or another and gave invaluable advice. Mary Fran Sowers, Hugh Joseph, Eric Rimm, and Barbara Underwood reviewed specific chapters and provided important comments. Jeffrey House at Oxford University Press gave essential support and counsel during the production of the book.

Contents

Nutritional Epidemiology

1

Overview of Nutritional Epidemiology

The field of nutritional epidemiology has developed from interest in the concept that aspects of diet may influence the occurrence of human disease. Although it is relatively new as a formal area of research, investigators have used basic epidemiologic methods for more than 200 years to identify numerous essential nutrients. In the mid-eighteenth century, observations that fresh fruits and vegetables could cure scurvy led Lind (1753) to conduct one of the earliest controlled clinical trials; lemons and oranges had the "most sudden and good effects" on the course of this disease, which was ultimately found to be the result of vitamin C deficiency. In an example from the late nineteenth century, the unusual occurrence of beri-beri among sailors subsisting largely on polished rice led Takaki to hypothesize that some factor was lacking in their diet; the addition of milk and vegetables to their rations effectively eliminated this disease. Decades later a deficiency of thiamine was found to be primarily responsible for this syndrome (Davidson and Passmore, 1971). Similarly, Goldberger used epidemiologic methods to determine that pellagra was a disease of nutritional deficiency, primarily associated with a corn meal subsistence diet in the southern United States (Terris, 1963). More recently, Chinese investigators determined by epidemiologic means that selenium deficiency is responsible for the high incidence of Keshan disease in central China (Guang-Qi, 1987).

Typically, deficiency syndromes occur with high frequency among those with very low intake and rarely or never occur among those not so exposed. In addition, these deficiency diseases often have short latent periods; symptoms are usually manifested within months of starting a deficient diet and can typically be reversed within days or weeks. Hence, research has moved rapidly from observations to experiments in both animals and humans.

Deficiency states for essential nutrients, such as scurvy and rickets, differ from most issues confronting nutritional epidemiologists today. The primary focus of contemporary nutritional epidemiology has been the major diseases of Western civilization, particularly heart disease and cancer. More recently, osteo-

porosis, cataracts, stroke, diabetes, and congenital malformations also have been the object of such research. Unlike nutritional deficiencies, these diseases almost certainly have multiple causes, potentially including not only diet but also genetic, occupational, psychosocial, and infectious factors; level of physical activity; behavioral characteristics such as cigarette use; and other influences. These multiple potential determinants may act alone or in combination. Also, many of these diseases have long latent periods; they may sometimes result from cumulative exposure over many years and in other instances from a relatively short exposure occurring many years before diagnosis. For most of these diseases, the relevant period of exposure is unknown. It is possible that exposures with short latent periods are also important for some of these diseases. For example, it is conceivable that smoking a cigarette or consuming large amounts of a certain food could, within hours, precipitate an acute myocardial infarction or thrombotic stroke by altering blood coagulability even though the underlying atherosclerosis has accumulated over decades. A third characteristic of these diseases is that they occur with relatively low frequency despite a substantial cumulative lifetime risk. In addition, these conditions are not readily reversible, and may result from excessive, as well as insufficient, intake of dietary factors. All these features have important implications for the design of studies to elucidate their etiologies.

The relationships between diet and the occurrence of the major diseases of our civilization are of both scientific and practical importance to public health nutritionists. The classic methods of nutritionists, such as basic biochemistry, animal experimentation, and metabolic studies in humans, contribute substantially but do not directly address these relationships. These issues should fall naturally within the realm of epidemiology, a discipline whose focus is the occurrence of human disease. Although epidemiologic efforts originally concentrated primarily on infectious diseases, during the last 30 years attention has largely shifted to the etiology of chronic diseases. Thus contemporary epidemiologists are accustomed to the study of diseases with low frequencies, long latency periods, and multiple causes. For example, hypertension, hypercholesterolemia, and cigarette smoking have been identified as major determinants of coronary heart disease; this knowledge has contributed to a major decline in this cause of death during recent years.

Although epidemiology is logically equipped to address the dietary causes of disease, the complex nature of diet poses an unusually difficult challenge to this discipline (Willett, 1987). Cigarette smoking is more typical of exposures studied by epidemiologists: subjects or their spouses can report whether or not they smoke cigarettes with a high degree of accuracy. Furthermore, individuals can readily provide quantitative information on the number of cigarettes they smoke per day, their usual brand of cigarettes, the age at which they started smoking, and changes in their pattern of use. The ease with which relatively accurate information on cigarette smoking can be obtained has contributed to the rapid accumulation of an enormous and remarkably consistent literature on the health effects of this habit. Diet, in contrast, represents an unusually complex set of exposures that are strongly intercorrelated. With few exceptions, all indi-

viduals are exposed to hypothesized causal factors; everyone eats fat, fiber, and vitamin A, for instance. Thus, exposures cannot be characterized as present or absent; rather they are continuous variables, often with a rather limited range of variation. Furthermore, individuals rarely make clear changes in their diet at identifiable points in time; more typically, eating patterns evolve over periods of years. Finally, individuals are generally not aware of the content of the foods that they eat; therefore, the consumption of nutrients is usually determined indirectly based on the reported use of foods or on the level of biochemical measurements.

Thus, the most serious limitation to research in nutritional epidemiology has been the lack of practical methods to measure diet. Because such epidemiologic studies usually involve at least several hundred and sometimes tens of thousands of subjects, dietary assessment methods must be not only reasonably accurate, but also relatively inexpensive.

The difficulties in assessing diet have led some epidemiologists to believe that it is unlikely that useful measurements of the diets of individual subjects within free-living populations can be collected at all (Wynder, 1976). In addition, some believe that the diets of persons within one country are too homogeneous to detect relationships with disease (Hebert and Wynder, 1987). Much of this skepticism arose from the intense interest in dietary lipids, serum cholesterol, and coronary heart disease beginning in the 1950s. Although it has been demonstrated in controlled metabolic studies that increases in saturated fat and cholesterol or decreases in polyunsaturated fat raise serum cholesterol, no correlation between intake of these lipids and serum cholesterol was found in many cross-sectional studies within the United States (Jacobs et al., 1979). Many concluded that as an association clearly did exist, the measurement of diet was so inaccurate that the relationship was obscured. In retrospect, it is apparent that any expectation of a strong correlation is unrealistic, and that a lack of correlation has several explanations. Most importantly, serum cholesterol is relatively insensitive to dietary lipid intake; metabolic studies clearly show that substantial changes in dietary intake produce rather modest changes in serum cholesterol (see Chapter 12). The expected correlation between cholesterol intake and serum cholesterol in the populations studied would be only on the order of 0.10, even with a perfect measure of dietary intake, because many factors, including genetic determinants, influence serum cholesterol.

Example: In the data of Shekelle et al. (1981), the standard deviation (SD), for dietary cholesterol is 68 mg/1000 kcal and the standard deviation for serum cholesterol is 54 mg/dl. From metabolic ward studies of Mattson et al. (1972), a 10-mg/1000-kcal change in dietary cholesterol causes 1.2 mg/dl change in serum cholesterol; thus the expected SD in serum cholesterol due to dietary cholesterol variation is 8.2 mg/dl. The theoretically expected correlation between cholesterol intake and serum cholesterol is

$$r = \frac{\text{Expected SD due to diet}}{\text{Total SD for serum cholesterol}} = \frac{8.2}{54} = 0.15$$

The standard deviation in this example for dietary cholesterol is, in reality, overstated as it included measurement error as well as true variation. Indeed, Shekelle et al. (1981), found that the relationship (using regression analysis) between dietary lipid intake based on carefully conducted interviews and serum cholesterol was not markedly different than that obtained from metabolic studies. Because of other determinants of serum cholesterol, however, the correlation between dietary lipid intake and serum cholesterol was only 0.08; as the study population was large, this small correlation was highly statistically significant. Furthermore, some factors are associated with both reduced cholesterol intake and higher serum cholesterol, such as low levels of physical activity and knowledge of hypercholesterolemia. These tend to distort the true relationship between diet and serum levels toward an inverse association in cross-sectional studies (Shekelle et al., 1982). In studies that examined the relationships between *change* in diet and *change* in serum cholesterol using simple questionnaires, it has been possible to demonstrate rather strong correlations (see Chapter 6). The methods of dietary intake measurement in many of the early studies were very inadequate; most commonly these investigations used 24-hour recall methods, which provide a poor assessment of usual intake (see Chapter 3). Nevertheless, it is clear that the use of serum cholesterol as a criterion for the validity of a dietary measurement method is inappropriate. From the standpoint of the development of the field of nutritional epidemiology, the early focus on serum cholesterol was unfortunate. If beta-carotene had been of major interest at that time, the attitude toward the possibility of measuring diet in epidemiologic studies might have been different, as it is easy to demonstrate an association between intake measured by simple questionnaires and blood levels (see Chapter 6).

The resurgent interest in dietary etiologies of disease, particularly cancer, has stimulated the development and evaluation of methods for dietary assessment in epidemiologic applications. Many, although not all, aspects of diet can now be measured readily and inexpensively with sufficient accuracy to provide useful information. These methods, consisting of food intake, biochemical and anthropometric measurements, are discussed in detail in later chapters. Equally important, biologically meaningful between-person variation has been documented to exist within populations for most nutrients that have been studied; without such variation, observational studies of individuals would not be feasible.

EPIDEMIOLOGIC APPROACHES TO DIET AND DISEASE

The concepts, hypotheses, and techniques of nutritional epidemiology are derived from many sources. Biochemistry, for example, has provided findings that certain nutrients function as antioxidants that may protect critical cell components from damage, potentially reducing the incidence of cancer (Ames, 1983). Cell culture methods have been used to identify compounds, such as preformed vitamin A, that regulate growth and differentiation of cells, and that may, therefore, influence the risk of cancer in humans. Experiments in laboratory animals have provided much information regarding the effects of diet on the occurrence of disease and mechanisms of action. Metabolic and biochemical

studies among human subjects have yielded essential information on the physiologic effects of dietary factors. Findings from in vitro studies and animal experiments, however, cannot be extrapolated directly to humans (Ames et al., 1987), and physiologic and metabolic changes are several steps removed from the actual occurrence of disease in humans; thus, epidemiologic approaches are needed to address diet and disease relationships directly. Nevertheless, these basic science areas provide critical direction for epidemiologists and important information for the interpretation of their findings.

Correlation Studies

Until recently, epidemiologic investigations of diet and disease consisted largely of ecological or correlational studies, that is, comparisons of disease rates in populations with the population per capita consumption of specific dietary factors. Usually the dietary information in such studies is based on *disappearance data*, meaning the national figures for food produced and imported minus the food that is exported, fed to animals, or otherwise not available for humans. Many of the correlations based on such information are remarkably strong; for example, the correlation between meat intake and incidence of colon cancer is 0.85 for men and 0.89 for women (Armstrong and Doll, 1975; Fig. 1–1).

The use of international correlational studies to evaluate the relationships between diet and disease has several strengths. Most importantly, the contrasts in dietary intake are typically very large. For example, within the United States, most individuals consume between 30 and 45 percent of their calories from fat, whereas the *mean* fat intake for populations in different countries varies from

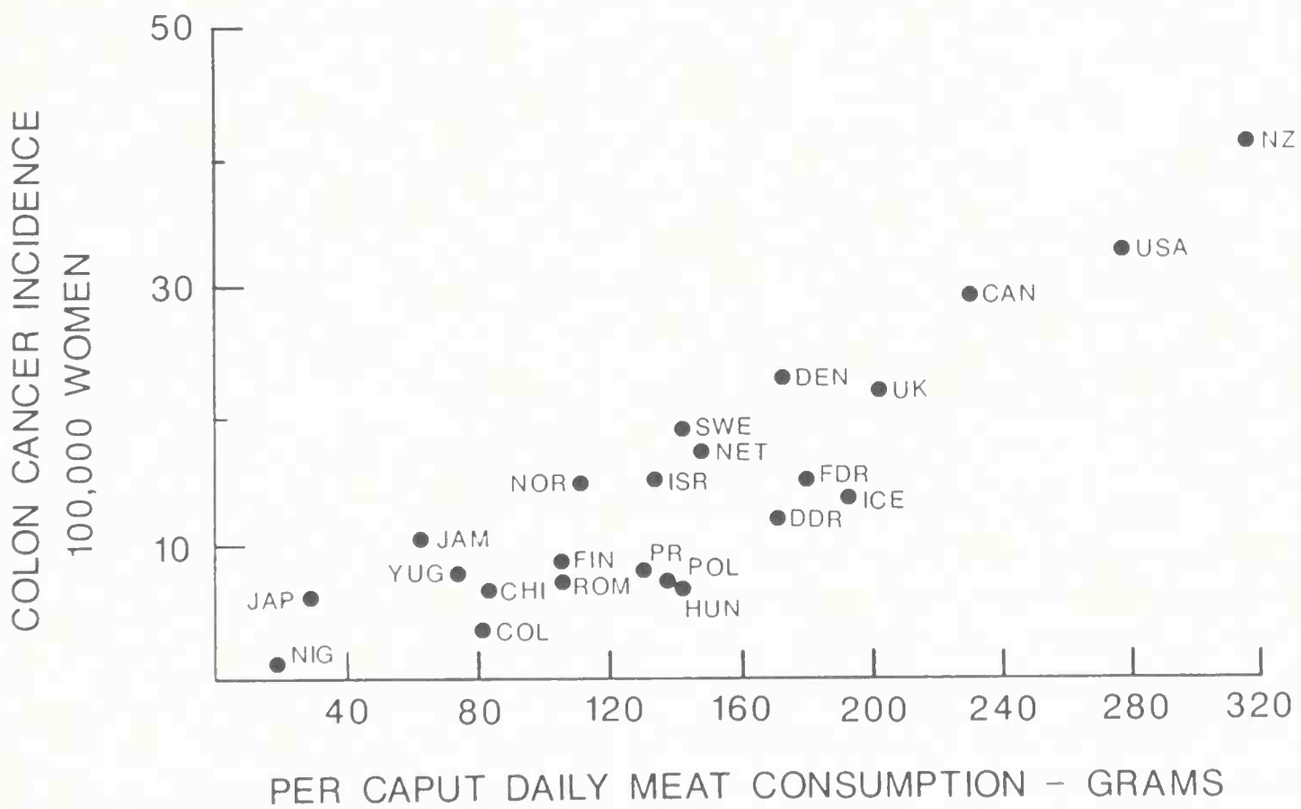


Figure 1–1. Correlation between per capita meat intake and incidence of colon cancer in women in 23 countries. (From Armstrong and Doll, 1975, reproduced with permission.)

11 to 42 percent of calories (Hebert and Wynder, 1987). Second, the average of diets for persons residing in a country are likely to be more stable over time than are the diets of individual persons within the country; for most countries the changes in per capita dietary intakes over a decade or two are relatively small. Finally, the cancer rates on which international studies are based are usually derived from relatively large populations and are, therefore, subject to only small random errors.

The primary problem of such correlational studies is that many potential determinants of disease other than the dietary factor under consideration may vary between areas with a high and low incidence of disease. Such confounding factors can include genetic predisposition; other dietary factors, including the availability of total energy intake; and other environmental or life-style practices. For example, with few exceptions, such as Japan, countries with a low incidence of colon cancer tend to be economically underdeveloped. Therefore, any variable related to industrialization will be similarly correlated with the incidence of colon cancer. Indeed, the correlation between gross national product and colon cancer mortality rate is 0.77 for men and 0.69 for women (Armstrong and Doll, 1975). More complex analyses can be conducted of such ecologic data that control for some of the potentially confounding factors. For example, McKeown-Eyssen and Bright-See (1985) have found that an inverse association of per capita dietary fiber intake and national colon cancer mortality rates persists after adjustment for fat intake.

Most correlational studies are also limited by the use of disappearance data that are only indirectly related to intake and are likely to be of variable quality. For example, the higher "disappearance" of calories per capita for the United States compared with most countries is probably related in part to wasted food in addition to higher actual intake. In addition, aggregate data for a geographic unit as a whole may be only weakly related to the diets of those individuals at risk of disease. As an extreme example, the interpretation of correlational data regarding alcohol intake and breast cancer is complicated as, in some cultures, most of the alcohol is consumed by men, but it is the women who develop breast cancer. These issues of data quality can potentially be addressed by collecting information on actual dietary intake in a uniform manner from the population subgroups of interest. This is currently being done in a study conducted in 65 geographic areas within China that are characterized by an unusually large variation in rates of many cancers (Chen et al., 1987).

Another serious limitation of the international correlational studies is that they cannot be independently reproduced, which is an important part of the scientific process. Although the dietary information can be improved and the analyses can be refined, the resulting data will really not be independent; the populations, their diets, and the confounding variables are the same. Thus, it is not likely that many new insights will be obtained from further ecologic studies among countries. For this reason, the methodologic aspects of correlational studies are not discussed further in this book.

The role of correlational studies in nutritional epidemiology is controversial. Clearly these analyses have stimulated much of the current research on diet and cancer and in particular they have emphasized the major differences in cancer

rates among countries. Traditionally, such studies have been considered the weakest form of evidence, primarily due to the potential for confounding by factors that are difficult to measure and control (Kinlen, 1983). More recently, some have felt that such studies provide the strongest form of evidence for evaluating hypotheses relating diet to cancer (Hebert and Wynder, 1987; Prentice et al., 1988). On balance, ecologic studies have unquestionably been useful, but are not sufficient to provide conclusions regarding the relationships between dietary factors and disease and may sometimes be completely misleading.

Special Exposure Groups

Groups within a population that consume unusual diets provide an additional opportunity to learn about the relation of dietary factors and disease. These groups are often defined by religious or ethnic characteristics and provide many of the same strengths as ecologic studies. In addition, the special populations often live in the same general environment as the comparison group, which may somewhat reduce the number of alternative explanations for any differences that might be observed. For example, the observation that colon cancer mortality in the largely vegetarian Seventh Day Adventists is only about half that expected (Phillips et al., 1980) has been used to support the hypothesis that meat consumption is a cause of colon cancer.

Findings based on special exposure groups are subject to many of the same limitations as ecologic studies. Many factors, both dietary and nondietary, are likely to distinguish these special groups from the comparison population. Thus, another possible explanation for the lower colon cancer incidence and mortality among the Seventh Day Adventist population is that differences in rates are attributable to a lower intake of alcohol or higher vegetable consumption. Given the many possible alternative explanations, such studies may be particularly useful when a hypothesis is *not* supported. For example, the finding that the breast cancer mortality rate among the Seventh Day Adventists is not appreciably different than the rate among the general United States population provides fairly strong evidence that meat eating does not cause a major increase in the risk of breast cancer (see Chapter 14).

Migrant Studies and Secular Trends

Migrant studies have been particularly useful in addressing the possibility that the correlations observed in the ecologic studies are due to genetic factors. For most cancers, populations migrating from an area with its own pattern of cancer incidence rates acquire rates characteristic of their new location (Staszewski and Haenszel, 1965; Adelstein et al., 1979; McMichael and Giles, 1988), although, for a few tumor sites, this change occurs only in later generations (Haenszel et al., 1972; Buell, 1973). Therefore, genetic factors cannot be primarily responsible for the large differences in cancer rates among these countries. Migrant studies are also useful to examine the latency or relevant time of exposure (see Chapter 14).

Major changes in the rates of a disease within a population over time provide

evidence that nongenetic factors play an important role in the etiology of that disease. In the United States, for example, rates of coronary heart disease rose dramatically over the first half of this century, and then subsequently declined (Working Group on Arteriosclerosis, 1981). These secular changes clearly demonstrate that environmental factors, possibly including diet, are primary causes of this disease, even though genetic factors may still influence who becomes affected given an adverse environment.

Case-Control and Cohort Studies

Many of the weaknesses of correlational studies are potentially avoidable in case-control studies (in which information about previous diet is obtained from diseased patients and compared with that from subjects without the disease) or cohort investigations (in which information on diet is obtained from disease-free subjects who are then followed to determine disease rates according to levels of dietary factors). In such studies, the confounding effects of other factors can be controlled either in the design (by matching subjects to be compared on the basis of known risk factors or by restriction), or in the analysis (by any of a variety of multivariate methods) if information has been collected on the confounding variables. Furthermore, dietary information can be obtained for the individuals actually affected by disease, rather than using the average intake of the population as a whole.

Case-control studies generally provide information more efficiently and rapidly than cohort studies as the number of subjects is typically far smaller and no follow-up is necessary. It remains unclear, however, whether consistently valid results can be obtained from case-control studies of dietary factors and disease because of the inherent potential for methodologic bias. This potential for bias is not unique for diet but is likely to be unusually serious for several reasons. Due to the limited range of variation in diet within most populations and some inevitable error in measuring intake, realistic relative risks in most studies of diet and disease are likely to be modest, say on the order of 0.5 to 2.0. These relative risks may seem small, but would be quite important because the prevalence of exposure is high. Given typical distributions of dietary intake, these relative risks are usually based on differences in means for cases and controls (or those who become cases and those who remain noncases in prospective studies) of only about 5 percent (see Chapters 3 and 12). It is thus obvious that a systematic error of even 3 or 4 percent can seriously distort such a relationship. In case-control studies it is easy to imagine that biases (due to selection or recall) of this magnitude could often occur, and it is extremely difficult to exclude the possibility that this degree of bias has not occurred in any specific study. Hence, it would not be surprising if case-control studies of dietary factors lead to inconsistent findings.

The selection of an appropriate control group for a study of diet and disease is also usually problematic. One common practice is to use patients with another disease as a control group, with the assumption that the exposure under study is unrelated to the condition of this control group. Because diet may well affect many diseases, it is often difficult to identify disease groups that are definitely

unrelated to the aspect of diet under investigation. A common alternative is to use a sample of persons from the general population as the control group. In many areas, particularly large cities, participation rates are low; it is now common for only 60 or 70 percent of eligible population controls to complete an interview (Hartge et al., 1984). Because diet is particularly associated with the level of general health consciousness, the diets of those who participate may differ substantially from those who do not; unfortunately little information is available that directly bears on this issue.

The many potential opportunities for methodologic bias in case-control studies of diet raise a concern that incorrect associations may frequently occur. Even if many studies arrive at correct conclusions, distortion of true associations in a substantial percentage produces an inconsistent body of published data, making a coherent synthesis difficult or impossible for a specific diet and disease relationship. Methodologic sources of inconsistency may be particularly troublesome in nutritional epidemiology due to the inherent biologic complexity of nutrient–nutrient interactions. As the effect of one nutrient may depend on the level of another (which can differ between studies and may not have even been measured), such interactions may result in apparently inconsistent findings in the context of epidemiologic studies. Thus, compounding biologic complexity with methodologic inconsistency may result in an uninterpretable literature. Existing data do not provide a clear answer as to whether consistent findings can be expected to accrue from case-control studies of diet. In studies of green and yellow vegetable intake in relation to lung cancer in men, remarkably consistent inverse associations have been found (see Chapter 13). On the other hand, findings from studies of fiber and fat intake in relation to large bowel cancer have been inconsistent (Willett and MacMahon, 1984).

Prospective cohort studies avoid most of the potential sources of methodologic bias associated with case-control investigations. Because the dietary information is collected before the diagnosis of disease, illness cannot affect the recall of diet. Distributions of dietary factors in the study population may be affected by selective participation in the cohort; however, low participation rates at enrollment will not distort the relationships between dietary factors and disease. Although losses to follow-up that vary by level of dietary factors can result in distorted associations in a cohort study, follow-up rates tend to be rather high as participants have already provided evidence of willingness to participate and they may also be followed passively by means of disease registries and vital record listings (Stampfer et al., 1984). In addition to being less susceptible to bias, prospective cohort studies provide the opportunity to obtain repeated assessments of diet over time and to examine the effects of diet on a wide variety of diseases, including total mortality, simultaneously.

The primary constraints on prospective studies of diet are practical. Even for common diseases, such as myocardial infarction or cancers of the lung, breast, or colon, it is necessary to enroll tens of thousands of subjects. The use of structured, self-administered questionnaires has made studies of this size possible, although still expensive. For diseases of somewhat lower frequency, however, even very large cohorts will not accumulate a sufficient number of cases within a reasonable amount of time. Case-control studies, therefore, continue to play

an important role in nutritional epidemiology. Given current uncertainty about measuring diets early in life, it is possible that neither study design will be able to address the influence of childhood diet on disease occurring decades later.

Controlled Trials

The most rigorous evaluation of a dietary hypothesis is the randomized trial, optimally conducted as a double-blind experiment. The principal strength of a randomized trial is that potentially distorting variables should be distributed at random between the treatment and control groups, thus minimizing the possibility of confounding by these extraneous factors. In addition, it is sometimes possible to create a larger contrast between the groups being compared by use of an active intervention. Such experiments among humans, however, are best justified after considerable nonexperimental data have been collected to ensure that benefit is reasonably probable and that an adverse outcome is unlikely. Experimental studies are particularly practical for evaluating hypotheses that minor components of the diet, such as trace elements or vitamins, can prevent disease as these nutrients can be formulated into pills or capsules (Stampfer et al., 1985).

Even if feasible, randomized trials of dietary factors and disease are likely to encounter several limitations. The time between change in the level of a dietary factor and any expected change in the incidence of disease is typically uncertain. Therefore, trials must be of long duration, and usually one cannot eliminate the possibility that any lack of difference between treatment groups may be due to insufficient duration. Compliance with the treatment diet is likely to decrease during an extended trial, particularly if treatment involves a real change in food intake, and the control group may well adopt the dietary behavior of the treatment group if the treatment diet is thought to be beneficial. Such trends, which were found in the Multiple Risk Factor Intervention Trial of coronary disease prevention (Multiple Risk Factor Intervention Trial Research Group, 1982), may obscure a real benefit of the treatment.

A related potential limitation of trials is that participants who enroll in such studies tend to be highly selected on the basis of health consciousness and motivation. It is possible, therefore, that the subjects at highest potential risk on the basis of their dietary intake, and thus susceptible to intervention, are seriously underrepresented. For example, if low beta-carotene intake is thought to be a risk factor for lung cancer and a trial of beta-carotene supplementation is conducted among a health conscious population that includes few individuals with low beta-carotene intake, one might see no effect simply because the study population was already receiving the maximal benefit of this nutrient through its usual diet. In such an instance, it would be useful to measure dietary intake of beta-carotene before starting the trial. Because the effect of supplementation is likely to be greatest among those with low dietary intakes, it would be possible either to exclude those with high intakes (the potentially nonsusceptibles) either before randomization or in subanalyses at the conclusion of the study. This requires, of course, a reasonable measurement of dietary intake.

It is sometimes said that trials provide a better quantitative measurement of the effect of an exposure or treatment because the difference in exposure between

groups is better measured than in an observational study. Although this contrast may at times be better defined in a trial (it is usually clouded by some degree of noncompliance), trials still usually produce an imprecise measure of the effect of exposure due to marginally adequate sample sizes and ethical considerations that require stopping soon after a statistically significant effect is seen. For example, with a p value close to 0.05, as was found in the Lipid Research Clinics Coronary Primary Prevention Trial (Lipid Research Clinics Program, 1984), the 95 percent confidence interval extends from no effect to a strong effect that is usually implausible. In an observational study an ethical imperative to stop does not exist when statistical significance occurs; continued accumulation of data can provide increasing precision regarding the relation between exposure and disease. A trial can provide unique information on the latent period between change in an exposure and change in disease; because spontaneous changes in diet are typically not clearly demarcated in time, the estimation of latent periods for dietary effects is usually difficult in observational studies.

Although all hypotheses would ideally be evaluated in randomized trials, this is sometimes impossible for practical or ethical reasons. For example, our knowledge of the effects of cigarette smoking on risk of lung cancer is based on observational studies, and it is similarly unlikely that randomized trials could be conducted to examine the effect of alcohol use on human breast cancer risk. It remains unclear whether trials of sufficient size, duration, and degree of compliance can be conducted to evaluate many hypotheses that involve major changes in eating patterns, such as a reduction in fat intake.

INTERPRETATION OF EPIDEMIOLOGIC DATA

The interpretation of positive (or inverse) associations in epidemiologic studies has received considerable attention; however, the evaluation of null or statistically nonsignificant findings has received less. Because either finding is potentially important, both are considered here.

If an association is observed in a epidemiologic study, we are usually concerned whether it represents a true cause-and-effect relationship, that is, if we actively changed the exposure, would that influence the frequency of disease? Hill (1965) has discussed factors that have frequently been considered as criteria for causality. These have included the strength of association, the consistency of a finding in various studies and populations, the presence of a dose-response gradient, the appropriate temporal relationship, the biologic plausibility, and the coherence with existing data. As has been pointed out by Rothman (1986), these cannot be considered as criteria as exceptions are likely to be frequent; this is particularly true in nutritional epidemiology. In this field, true associations are not likely to be strong although relative risks of 0.7 or 1.5 could potentially be important because the dietary exposures are common. The consistent finding of an association that cannot be explained by other factors in various populations markedly reduces the possibility that chance explains the findings and increases the likelihood of causality. Although the reproducibility of findings is extremely important, null findings should sometimes be expected in nutritional epidemi-

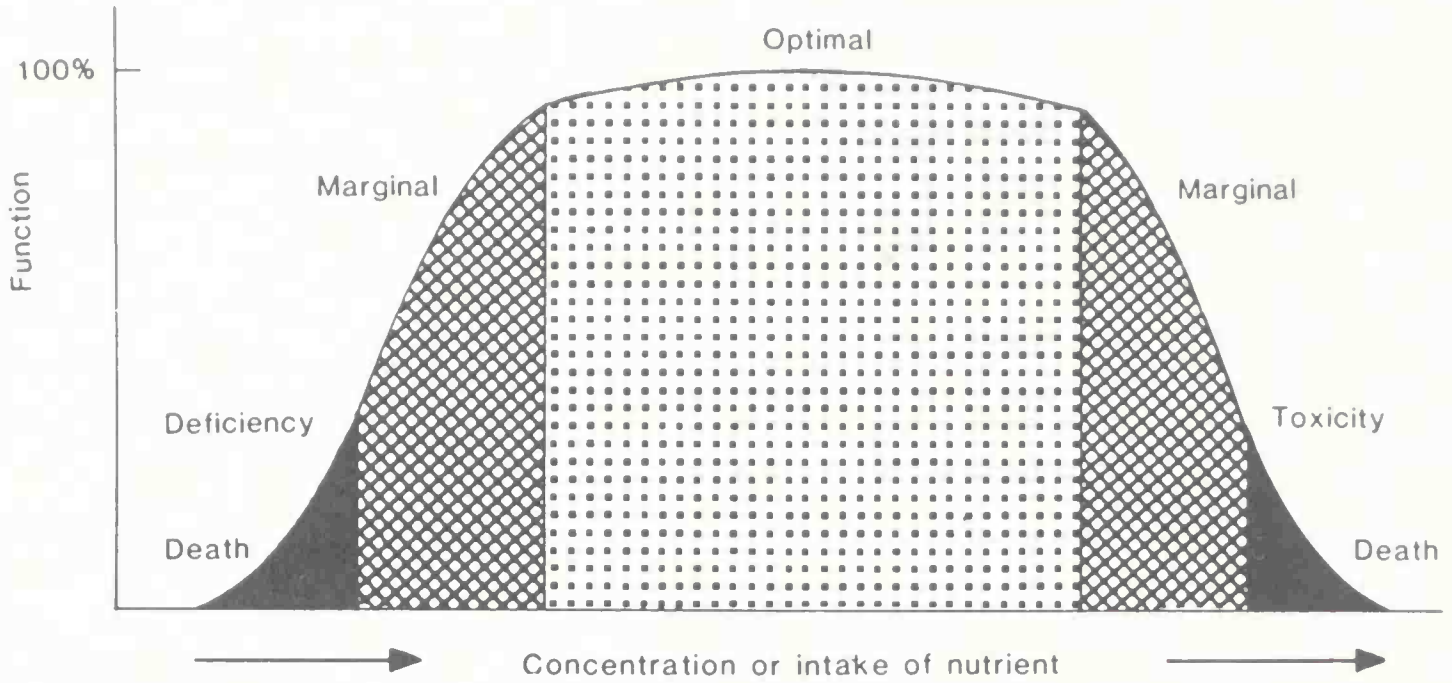


Figure 1–2. Hypothetical relationship between intake of an essential dietary factor and health (from Mertz, 1981, reproduced with permission). If two points on the ascending part of the curve are compared, it might be concluded that the nutrient was beneficial; if points on the horizontal portion were compared, it might be concluded that the nutrient had no effect; if points on the descending segment were contrasted, it might be reported that the nutrient was deleterious. The health effects of the nutrient can only be fully appreciated by an examination of the dose-response relationship over the full range of exposures, which may not be possible within any single study.

ology even when a causal relationship may exist; thus, absolute consistency is not a realistic expectation. Dose-response relationships are likely to be nonlinear and may be of almost any shape depending on our starting point on a hypothetical spectrum of exposure (Fig. 1–2). Moreover, apparently clear dose-response relationships can easily be the result of bias or confounding. Although compatibility of a finding with an established mechanism of disease causation supports causality, post hoc biologic explanations should be viewed cautiously as they can usually be developed for most observations, including those that are later refuted. Moreover, the pathophysiology of most cancers and many other chronic diseases is poorly understood, so that lack of a well-defined mechanism should not be construed as evidence against causality.

Knowledge that an association exists, even if deemed causal, is not sufficient to make public or personal decisions. Such actions require some knowledge of the shape and quantitative aspects of the dose-response relationship. For instance, knowledge that total fat intake is associated with risk of colon cancer would not provide a sufficient basis to recommend a universal reduction in fat intake. It would be much more useful to know, for example, the change in risk associated with a decrease in fat intake from 40 to 30 percent of total energy intake, which has been considered realistic for the U.S. population (Committee on Diet, Nutrition and Cancer, 1982), as well as the effect of a change from 30 to 20 percent of calories, which probably represents a limit of feasibility for the United States. It is entirely possible that a strong relation between fat intake and colon cancer risk exists below 20 percent of calories, but that above that level the relationship is nonlinear, flat, or too weak to be of importance (McMichael

and Potter, 1985). In addition to this information, knowledge of the approximate latent period between alteration in diet and change in disease incidence would be important. If this were several decades, older individuals might rationally ignore the association in making decisions regarding their diet.

Interpretation of Null Associations

In a study of diet and disease, failure to observe a statistically significant association when such an association truly exists can occur in several circumstances alluded to earlier. One possibility is that the variation in diet is insufficient; in the extreme, no associations can occur if everyone in the study population eats the same diet. Second, variation may exist for the study population, but only within a "flat" portion of the total dose-response relationship. A third possibility is that the method of measuring dietary intake is not sufficiently precise to measure differences that truly exist. Fourth, an association may be missed because of low statistical power due to an inadequate number of diseased and nondiseased subjects. Fifth, a relationship could be undetected because the temporal relationship between the measured exposure and the occurrence of disease did not encompass the true latent period; this could easily happen if the critical dietary exposure occurred during childhood and the disease was diagnosed during adulthood. Sixth, an association could be undetected because some unmeasured third variable was related to exposure and disease in opposite directions, in other words, negative confounding existed. In addition to these six largely biologic reasons for failure to observe an association, methodologic sources of bias could obscure a relationship.

It is obviously not informative to describe a study as null or nonsignificant unless the possible explanations noted previously have been addressed. Clearly, no single study can fully encompass the total possible range of human diets, measure all aspects of diet with absolute precision, assess all potential latent periods, and control for all potentially confounding variables. What must be done, then, is to describe the conditions and limitations of the null findings. First, it is critical to demonstrate that true variation in diet exists within the study population and that the method of measuring diet provides useful discrimination among subjects (see Chapter 6). It is not adequate to demonstrate that dietary variation exists on the basis of measurements using the study instrument alone as this variation could merely represent error. On the other hand, demonstration that measurements made using the study instrument correlate with measurements made using another method with independent sources of error provides evidence both that diet does vary within the population and that the study instrument is capable of detecting this variation.

Although confidence intervals are important for reporting positive associations, they are even more critical for results that are near the null (e.g., relative risk = 1.0) or not statistically significant as they provide a sense of the range of values that are still consistent with the data. Although it has seldom been done in practice, confidence limits should ideally be adjusted for measurement error; measurement error tends to make the true confidence intervals wider than those usually calculated, assuming no such error (Rosner et al., 1988). It has become

fashionable, and even required by some editors, to include a priori power calculations in reports of study results. Because confidence intervals are determined by the observed data as well as the influence of chance, a priori power estimates add little once the study is completed. (The use of power estimates to interpret nonsignificant findings can easily be misleading; it is quite possible for a study to have low a priori power to detect a positive association but have confidence limits that widely exclude that positive association if the association is in the opposite direction.) The range of latent periods encompassed by the study should also be described; in dietary studies usually this is possible only crudely. If the study is a prospective cohort, or data are available in retrospect for several points in time, analyses can be conducted to evaluate associations separately for different latent periods (Rothman, 1986). Finally, it will be important to describe the dietary and nondietary correlates of the primary exposure that have been evaluated as potential confounding variables.

Because it is rare that all aspects of a hypothesis can be addressed in one study, it is important to describe which aspects have or have not been evaluated. For example, it is of limited use to conclude simply that a given study of dietary vitamin C intake and colon cancer was negative. It would be much more informative to say, for example, "Vitamin C intake determined by a detailed quantitative method was 40 mg per day for the tenth percentile and 200 mg per day for the 90th percentile. During a 5-year follow-up period the observed relative risk was 1.0 with a 95 percent confidence interval of 0.8 to 1.3 after adjusting for exposure measurement error for a difference of 50 mg per day of vitamin C intake, which corresponds to a 50 percent increase for the average subject. Finally, adjustment for parental history of colon cancer and intakes of dietary fiber and calcium did not alter the findings." It is thus clear from this description that the effects of very low and very high vitamin C intakes are not being evaluated nor is the influence of childhood diet, and that a 10 percent, but not a 30 percent, reduction in risk by a 50 percent increase in intake later in life is still quite possible.

Multivariate Relationships of Diet and Disease

Relationships between dietary factors and disease are likely to be extremely complex for both biologic and behavioral reasons. Types and amounts of food eaten may be related to important nondietary determinants of disease, such as age, smoking, exercise, and occupation, which may both distort or confound and modify relationships with diet. As discussed in the next chapter, intakes of specific nutrients tend to be intercorrelated so that associations with one nutrient may be confounded by other aspects of the diet. Furthermore, the intake of one nutrient may modify the absorption, metabolism, or requirement for another nutrient, thus creating a biologic interaction. Due to these complexities, it is generally unsatisfactory to examine the relationship between a single dietary factor and disease in isolation. In practice, it is almost always necessary to employ multivariate techniques, including both stratified analyses and statistical models, to adjust for potentially confounding variables and examine interactions. Nevertheless, this book focuses primarily on defining the association of one or

two dietary factors at a time, as more complex analyses must proceed from the capacity to measure clearly and interpret bivariate relationships. Moreover, the mechanical aspects of multivariate analysis are covered in other advanced epidemiology texts and are not unique to nutritional epidemiology.

The use of multivariate methods in any particular analysis requires a careful consideration of the precise question that is being posed and whether potential co-variates are true confounders as opposed to effects of the primary exposure. Confusion resulting from the inappropriate application of multivariate methods is illustrated by the controversy surrounding the relation of relative weight and risk of coronary heart disease (Manson et al., 1987). In a number of reports, blood pressure, glucose tolerance, and serum lipid levels were included in multivariate models along with a measure of relative weight. Because these other risk factors are strongly influenced by obesity and are thus in the causal pathway relating relative weight with coronary heart disease, their inclusion substantially diminishes the apparent effect of relative weight. Conclusions based on such analyses that obesity has little relationship with coronary heart disease are misleading as obesity cannot be stripped of its metabolic consequences by sophisticated statistical methods. The application of multivariate methods in nutritional epidemiology necessitates maximal use of existing knowledge regarding the effects of dietary factors to avoid similar problems in the future.

SUMMARY

Our knowledge is still largely incomplete regarding the relationship between dietary factors and the major illnesses of our culture. These illnesses include not only cancer and heart disease, which have received the most attention, but also congenital malformations, degenerative conditions of the eye, fractures, and many infectious diseases that are hypothesized to be influenced by the nutritional status of the host. Randomized trials may eventually provide definitive answers to some of these questions. Our knowledge of many of these relationships, however, depends largely on observational epidemiologic data for the near future and, for some relationships, indefinitely. For this reason, it is crucial to refine maximally our methods of data collection, analytic procedures, and interpretation of findings. The ensuing chapters are intended to further our progress in this direction.

REFERENCES

Adelstein, A. M., J. Staszewski, and C. S. Muir (1979). Cancer mortality in 1970–1972 among Polish-born migrants to England and Wales. *Br. J. Cancer 40*, 464–475.

Ames, B. N. (1983). Dietary carcinogens and anticarcinogens. *Science 221*, 1256–1264.

Ames, B. N., R. Magaw, and L. S. Gold (1987). Ranking possible carcinogenic hazards. *Science 236*, 271–280.

Armstrong, B. and R. Doll (1975). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *Int. J. Cancer 15*, 617–631.

Buell P. (1973). Changing incidence of breast cancer in Japanese–American women. *J.N.C.I. 51,* 1479–1483.

Chen, J., T. C. Campbell, L. Junyao and R. Peto (1987). The dietary, lifestyle, and mortality characteristics of 65 rural populations in the Peoples Republic of China (Monograph), Cornell Univ., Dept. of Nutrition.

Committee on Diet, Nutrition and Cancer (1982). *Diet, Nutrition and Cancer.* Washington, D.C.: National Academy Press.

Davidson, S. and R. Passmore (1971). *Human Nutrition and Dietetics.* Edinburgh and London: E. S. Livingstone.

Guang-Qi, Y. (1987). Research on selenium-related problems in human health in China. In Combs, G. F., J. E. Spallholz, O. R. Levander, and J. E. Oldfield, eds., *Selenium in Biology and Medicine,* Part A. New York: AVI Book, pp. 9–32.

Haenszel, W., M. Kurihara, M. Seig, and R.K.C. Lee (1972). Stomach cancer among Japanese in Hawaii. *J.N.C.I. 49,* 969–988.

Hartge, P., L. A. Brinton, J. F. Rosenthal, J. I. Cahill, R. N. Hoover, and J. Waksberg (1984). Random digit dialing in selecting a population-based control group. *Am. J. Epidemiol. 120,* 825–833.

Hebert, J. R. and E. L. Wynder (1987). Dietary fat and the risk of breast cancer. (letter) *N. Engl. J. Med 317,* 165.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proc. R. Soc. Medicine 58,* 295–300.

Jacobs, D. R., J. T. Anderson, and H. Blackburn (1979). Diet and serum cholesterol: Do zero order correlations negate the relationship? *Am. J. Epidemiol. 110,* 77–87.

Kinlen, L. (1983). Fat and cancer. *Br. Med. J. 286,* 1081–1082.

Lind, J. (1753). *Treatise of the Scurvy.* Reprinted by Edinburgh: University Press, 1953.

Lipid Research Clinics Program (1984). The Lipid Research Clinics Coronary Primary Prevention Trial Results. I. Reduction in incidence of coronary heart disease. *J.A.M.A. 251,* 351–364.

Manson, J. E., M. J. Stampfer, C. H. Hennekens, and W. C. Willett (1987). Body weight and longevity. *J.A.M.A. 257,* 353–358.

Mattson, E. H., B. A. Erickson, and A. M. Klegman (1972). Effect of dietary cholesterol on serum cholesterol in man. *Am. J. Clin. Nutr. 25,* 589–594.

McKeown-Eyssen, E. and E. Bright-See (1985). Dietary factors in colon cancer: International relationships. An update. *Nutr. Cancer 7,* 251–253.

McMichael, A. J. and J. D. Potter (1985). Diet and colon cancer: Integration of the descriptive, analytic, and metabolic epidemiology. *Natl. Cancer Inst. Monogr. 69,* 223–228.

McMichael, A. J. and G. G. Giles (1988). Cancer in migrants to Australia: Extending the descriptive epidemiologic data. *Cancer Res 48,* 751–756.

Mertz, W. (1981). The essential trace elements. *Science, 213,* 1332–1338.

Multiple Risk Factor Intervention Trial Research Group (1982). Multiple Risk Factor Intervention Trial. Risk Factor Changes and Mortality Results. *J. Am. Med. Assoc. 248,* 1465–1477.

Phillips, R. L., L. Garfinkel, J. W. Kuzma, W. L. Beeson, T. Lotz, and B. Brin (1980). Mortality among California Seventh-Day Adventists for selected cancer sites. *J.N.C.I. 65,* 1097–1107.

Prentice, R. L., F. Kakar, S. Hursting, L. Sheppard, R. Klein, and L. Kushi (1988). Aspects of the rationale for the Women's Health Trial. *J.N.C.I. 80,* 802–814.

Rosner, B. A., W. C. Willett, and D. Spiegelman (1989). Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat. in Med.,* in press.

Rothman, K. J. (1986). *Modern Epidemiology.* Boston/Toronto: Little, Brown.

Shekelle, R. B., A. M. Shryock, et al. (1981). Diet, serum cholesterol, and death from coronary heart disease: The Western Electric Study. *N. Engl. J. Med. 304,* 65–70.

Shekelle, R. B., J. Stamler, O. Paul, A. M. Shryock, S. Liu, and M. Lepper (1982). Dietary lipids and serum cholesterol level: Change in diet confounds the cross-sectional association. *Am. J. Epidemiol. 115,* 506–514.

Stampfer, M. J., W. C. Willett, F. E. Speizer, D. C. Dysert, R. J. Lipnick, B. Rosner, and C. H. Hennekens (1984). Test of the National Death Index. *Am. J. Epidemiol. 119,* 837–839.

Stampfer, M. J., J. E. Buring, W. C. Willett, B. Rosner, K. Eberlein, and C. H. Hennekens (1985). The 2 x 2 factorial design: Its application to a randomized trial of aspirin and carotene in U.S. physicians. *Stat. in Med. 4,* 111–116.

Staszewski, J. and W. Haenszel (1965). Cancer mortality among the Polish-born in the United States. *J.N.C.I. 35,* 291–297.

Terris, M. (1964). Goldberger on pellegra. Baton Rouge Louisiana State Univ. Press.

Willett, W. C., and B. MacMahon (1984). Diet and Cancer—An Overview. *N. Engl. J. Med. 310,* 633–638 and 677–701.

Willett, W. (1987). Nutritional epidemiology: Issues and challenges. *Int. J. Epidemiol. 16* (Suppl.) 312–317.

Working Group on Arteriosclerosis of the National Heart, Lung, and Blood Institute (1981). *Arteriosclerosis 1981 Vol. 2.* U.S. Dept. HHS, NIH Publication No. 82-2035.

Wynder E. L. (1976). Nutrition and cancer. *Fed. Proc. 35,* 1309–1315.

2

Foods and Nutrients

The complexity of the human diet represents a daunting challenge to anyone contemplating a study of its relation to disease. The foods we consume each day contain literally thousands of specific chemicals, some known and well quantified, some characterized only poorly, and others completely undescribed and presently unmeasurable. The chemicals that comprise our food can be described by the following nonmutually exclusive categories:

1. *Essential nutrients.* The essential nutrients, which include minerals, vitamins, lipids, and amino acids, have little in common except that insufficient intake results in predictable clinical signs and symptoms of deficiency. Although it is likely that additional essential micronutrients remain to be identified, most are already well characterized; this represents the remarkable achievement of twentieth-century nutritional scientists.
2. *Major energy sources.* The substantial majority of the food we eat consists of proteins, carbohydrates, fats, and alcohol that are oxidized to provide energy. Proteins, carbohydrates, and fats are, of course, very heterogeneous and it seems likely that the mix of these fuels influences the long-term function of the human organism.
3. *Additives.* These substances are consciously added to our food for purposes such as preservation (e.g., nitrates, butylated hydroxy-toluene, and salt), coloring, and enhancement of consistency or flavor. Although such additives have elicited great public concern, they represent only a small fraction of the substances in our food and are among the best characterized and most closely regulated. The health effects of these additives have not been completely studied epidemiologically; however, evidence does not presently exist that they contribute importantly to disease in the United States (Doll and Peto, 1981).
4. *Agricultural chemical contaminants.* These products include pesticides, herbicides, fungicides, and growth hormones for both plants and animals.

5. *Microbial toxin contamination.* Aflatoxins produced by the mold *Aspergillus flavus* are a classic example of this class of substances. The contamination of grains and other plant products by this mold is widespread, particularly where storage conditions are poor. Although aflatoxins are likely to contribute to the high rates of liver cancer in many developing countries (Busby and Wogan, 1984), there is presently little direct evidence that they are responsible for disease in the industrialized countries.
6. *Inorganic contaminants.* A wide variety of other chemicals inadvertently enter our food supply, including metals such as cadmium and lead, and synthetic compounds such as polychlorinated biphenyls.
7. *Chemicals formed in the cooking or processing of food.* Burning or charring food creates many substances that are mutagenic using in vitro testing systems. Recently, it has been found that even heating meat products without burning can create a series of substances that are mutagenic (Ames et al., 1987; Sugimura, 1986). The relevance of these substances to human health remains unknown.
8. *Natural toxins.* Many plants have, through evolution, developed the capacity to produce chemicals that are toxic to the insects or other animals that might attack them. Although many plants are recognized as poisonous to humans, the foods we eat also contain natural pesticides even though they do not consistently produce acute symptoms.
9. *Other natural compounds.* In the process of eating plant and animal products, we consume the countless substances that are vital for maintaining the structure and function of these living cells. For example, we consume DNA and RNA, the specialized lipids of cell membranes, and thousands of enzymes and enzyme inhibitors. Although some of these compounds provide essential nutrients for humans, most are thought of as incidental to the human diet. It is likely, however, that many of these compounds influence the occurrence of chronic human diseases. For example, cholesterol has an important structural role in the membranes of animal tissues; however, the consumption of excessive cholesterol in the form of animal products is likely to contribute to the occurrence of coronary heart disease. More recently, accumulating evidence suggests that protease inhibitors contained in certain plant products, particularly legumes, may reduce the risk of some cancers (Troll et al., 1984).

The complex array of dietary chemicals outlined previously creates obvious challenges for investigators seeking to understand the relationships of diet with human disease. In particular, it is difficult to know where to begin the search and how to set priorities for investigation along the way. Up to this point, most research has focused on the first two categories: the essential nutrients and major energy sources. This has some justification as a starting point. In general, living organisms do not respond in either a linear or an all-or-nothing manner to increasing levels of an external exposure. This principle has been depicted graphically for essential nutrients by Mertz (Mertz, 1981, see Fig. 1–2). Given that low intake of an essential nutrient produces clinical dysfunction in the short run, it is reasonable to hypothesize that less extreme levels of intake might produce

subclinical dysfunction that affects the probability of developing chronic disease over a period of decades.

The additional focus on the major energy sources has justification in the simple fact that they are quantitatively so important in our diets. Furthermore, gross and obvious differences exist in the mix of these energy sources among human populations. As noted in Chapter 1, these differences are correlated with striking variation in rates of many diseases; the temptation to invoke a causal interpretation of these associations has been strong. Tradition based on the focus of nutritionists earlier in this century, and convenience have also probably contributed to the research emphasis on essential nutrients and energy sources. Many of us have become accustomed to first thinking of nutrients when the topic of diet is raised. Moreover, decades of work have been invested in improving available data on the nutrient content of foods, so that much better information is available for this group of dietary chemicals than for the other categories of substances in our foods.

Apart from essential nutrients and energy sources, the immense variety of other dietary substances confronts investigators with the need to establish research priorities. Additives and agricultural chemicals deserve attention because they may be controlled and because accidents can lead to high local concentrations. Ames and co-workers (1987) have provided an extensive review of dietary chemicals with respect to their potential for causing cancer. They pointed out that many of the naturally occurring substances are both more potent toxins or mutagens and far more abundant in food than the manmade chemicals that have been consciously or accidentally added to food. Although there is good reason for the research emphasis to date, it is important that tradition and convenience do not deter us from consideration that other aspects of diet may be important influences on the risk of human disease.

NUTRIENTS VERSUS FOODS

Throughout nutrition in general and in the preceding discussion, diet has usually been described in terms of its chemical composition, for example, its nutrient content. Alternatively, diet can be described in terms of foods or food groups. It may be useful to consider the advantages and disadvantages of representing diet in these ways.

The primary advantage of representing diets as specific compounds or groups of compounds is that such information can be directly related to our fundamental knowledge of biology. From a practical perspective, the exact structure of a compound must usually be known if it is to be synthesized and used for supplementation. In epidemiologic studies, calculation or measurement of total intake of a nutrient (as opposed to using the contribution of only one food at a time) provides the most powerful test of a hypothesis, particularly if a number of foods each contribute only modestly to intake of that nutrient. For example, in a particular study, it is quite possible that total fat intake could be clearly associated with risk of disease, whereas none of the contributions to fat intake by individual foods would, on their own, be significantly related to disease.

The use of foods to represent diet has several practical advantages when examining relationships with disease. Particularly when suspicion exists that some aspect of diet is associated with risk but a specific hypothesis has not been formulated, an examination of the relations of foods and food groups with risk of disease provides a means to explore the data. Associations observed with specific foods may lead to a hypothesis relating to a defined chemical substance. For example, observations that higher intakes of green and yellow vegetables were associated with reduced rates of lung cancer led to the hypothesis that beta-carotene might protect DNA from damage due to free radicals and singlet oxygen (Peto et al., 1981). The finding by Graham and co-workers (1978) that intake of cruciferous vegetables was inversely related to risk of colon cancer supported the suggestion that indole compounds contained in these vegetables may be protective (Wattenberg and Loub, 1978). Similarly, lower rates of coronary heart disease among Eskimos and individuals who consume higher amounts of fish has generated the hypothesis that omega-3 fatty acids reduce the risk of this disease, perhaps by inhibiting the formation of thromboxane and thus the propensity for intracoronary thrombus formation (Lands, 1986).

Even more seriously than the lack of a well-formulated hypothesis, the premature focus on a specific nutrient that turns out to have no relation with disease may lead to the erroneous conclusion that diet has no effect. Mertz (1984) has pointed out that foods are not fully represented by their nutrient composition, noting as an example that milk and yogurt produce different physiologic effects despite a similar nutrient content. Furthermore, the valid calculation of a nutrient intake from data on food consumption requires that reasonably accurate food composition information be available, which markedly constrains the scope of dietary chemicals that may be investigated, as such information exists for only several dozen commonly studied nutrients.

Epidemiologic analyses based on foods, as opposed to nutrients, are generally most directly related to dietary recommendations as individuals and institutions ultimately manipulate nutrient intake largely by their choice of foods. Even if the intake of a specific nutrient is convincingly shown to be related to risk of disease, this is not sufficient information on which to make dietary recommendations. Because foods are an extremely complex mixture of chemicals that may compete with, antagonize, or alter the bioavailability of any single nutrient contained in that food, it is not possible to predict with certainty the health effects of any food solely on the basis of its content of one specific factor. For example, there is concern that high intake of nitrates may be deleterious, particularly with respect to gastrointestinal cancer. The primary sources of nitrates in our diets, however, are green, leafy vegetables, which, if anything, appear to be associated with reduced risk of cancer at several sites. Similarly, because of the high cholesterol content of eggs, their avoidance has received particular attention in diets aimed at reducing the risk of coronary heart disease; per capita consumption of eggs declined by 25 percent in the United States between 1948 and 1980 (Welsh and Marston, 1982). Eggs, however, are more than cholesterol capsules; they provide a rich source of essential amino acids and micronutrients, and are relatively low in saturated fat. It is thus difficult to predict the net effect of egg consumption on risk of coronary heart disease, much less the effect on overall

health. At present there are few data indicating that individuals who consume more eggs have a higher risk of coronary heart disease.

Given the strengths and weaknesses of using nutrients or foods to represent diet, it appears that an optimal approach to epidemiologic analyses employs both. In this way, a potentially important finding is least likely to be missed. Moreover, the case for causality is strengthened when an association is observed with overall intake of a nutrient and also with more than one food source of that nutrient, particularly when the food sources are otherwise different. This provides, in some sense, multiple assessments of the potential for confounding by other nutrients; if an association was observed for only one food source of the nutrient, other factors contained in that food would tend to be similarly associated with disease. As an example, the hypothesis that alcohol intake causes breast cancer was strengthened by observing not only an overall association between alcohol intake and breast cancer risk, but also by independent associations between both beer and liquor intake and risk of breast cancer, thus making it less likely, but not impossible, that some factor other than alcohol in these beverages was responsible for the increased risk.

One practical drawback to the use of foods to represent diet is their large number and complex, often reciprocal, interrelationships that are largely due to individual behavioral patterns. For example, in one unpublished analysis, we found that potato chips was the food most strongly associated (inversely) with blood carotene levels, presumably because potato chip consumers tend to avoid vegetables and fruits in general. If carotene intake was truly protective for a certain disease, the resulting positive association between potato chip intake and disease could be misleading, as the relationship would have been indirect. Many other reciprocal relationships emerge upon perusal of typical datasets; for example, dark bread eaters tend not to eat white bread, margarine users tend not to eat butter, and skim milk users tend not to use whole milk. This complexity is, of course, one of the reasons to compute nutrient intakes that summarize the contributions of all foods.

An intermediate solution to the problem posed by the complex interrelationships among foods is to use food groups or to compute the contribution of nutrient intake from various food groups. For example, Manousos and co-workers (1983), combined the intakes of foods from several predefined groups to study the relation of diet with risk of colon cancer; they observed increased risk among subjects with high meat intake and with low consumption of vegetables. The computation of nutrient intakes from different food groups is illustrated by a prospective study among British bank clerks conducted by Morris and co-workers (1977), who observed an inverse relation between overall fiber intake and risk of coronary heart disease. It is well recognized that fiber is an extremely heterogeneous collection of substances and that the available food composition data for specific types of fiber is quite incomplete. Therefore, these authors computed fiber intake separately from various food groups and found that the entire protective effect was attributable to fiber from grains; fiber from fruits or vegetables was not associated with risk of disease. This analysis both circumvents the inadequacy of food composition databases, and provides information in a form that is directly useful to individuals faced with decisions regarding choices

of foods. Similarly, Witteman and colleagues (1987) found an inverse relation between overall intake of calcium and incidence of hypertension, as well as independent associations between calcium from dairy products, from grains, and from other sources. Because these various sources of calcium differed widely with respect to other factors, this added information provided further support for the overall finding.

Another approach for dealing with the complex intercorrelations of many foods is to use "pattern analysis" as suggested by Jacobson and Stanton (1986). This basically employs factor analysis or cluster analysis, which are both well-developed statistical techniques, to aggregate empirically individuals with similar diets based on their reported intake of foods, rather than according to predefined scales as is done in the calculation of nutrient intakes. These methods may be most useful when well-developed hypotheses do not exist or when associations with specific nutrients have not been found, possibly because they do not represent the relevant aspects of diet. A limitation of these approaches is the need to provide a post hoc interpretation for the factors or clusters that evolve from the data. Although their utility has not been established in nutritional research, such approaches deserve consideration, particularly in situations where our understanding of disease etiology is limited.

In summary, a single optimal representation of diet does not exist for epidemiologic analyses. The choices depend on the nature and refinement of the hypotheses being addressed, the availability of food composition data, and the form of the dietary data that are available. In general, maximal information is obtained when analyses are conducted at the levels of nutrients, foods, and food groups.

FOOD COMPOSITION DATA SOURCES AND COMPUTATION SYSTEMS

To calculate the total intake of a nutrient for each subject in a study, information is required on the content of each food that has been reported. In some instances, existing databases and computer software can be used, particularly for standard summaries from diet records or 24-hour recalls (see Chapter 4). In other situations, such as when a structured questionnaire has been created for a specific application, it is necessary to assemble a special database for this purpose. Because the size of many databases is formidable and these are frequently updated, it is not the objective of this section to provide primary information on food composition. Rather, existing options and considerations in selecting a source of data are discussed.

Before computing intake of a nutrient from information on the use of foods, it is important to consider whether such a calculation is appropriate. The fundamental assumption underlying such a calculation is that the nutrient content of a specific food is approximately constant; for example, that each carrot eaten by every subject has the same beta-carotene content. We know that this assumption is never completely correct; the beta-carotene content of a carrot will vary with the size of the carrot and degree of maturity. Moreover, plant geneticists

have selectively increased the carotene content of this vegetable in the United States because marketing experts have found that Americans prefer dark orange carrots. In other parts of the world, however, pale yellow carrots are especially valued and plants are selected for this attribute; thus substantial genetic variability contributes to differences in the carotenoid content of carrots. Further variation in the nutrient content of a food may arise from differences in the growing and harvesting conditions, processing, storage, and cooking. Although all these sources of variation contribute to error in the estimation of calculated nutrient intake, the real question is whether these sources of variation are quantitatively large enough to distort calculations seriously. In the instance of beta-carotene intake, the assumption of constant nutrient intake does not appear to be seriously violated; the finding that calculated beta-carotene intake is correlated with plasma beta-carotene level provides direct evidence that the calculations are providing informative data. Although the carotenoid content of carrots may vary by a factor of three or four, someone who eats carrots regularly will still, on average, consume more beta-carotene than someone who does not. Moreover, if long-term beta-carotene intake is of interest, much of the variation in nutrient composition will effectively be canceled out as we are really concerned whether the average beta-carotene content of carrots eaten by any one subject varies substantially from person to person. These differences in average beta-carotene content are probably substantially less than differences in beta-carotene content of carrots eaten by a person from one day to another. Such sources of variation in nutrient content of food appear not to have been addressed by any formal analysis.

Whereas the assumption of constant nutrient content in foods does not appear to be seriously violated in the case of beta-carotene, selenium provides an example where this assumption is sufficiently incorrect so as to preclude useful calculations of intake. The underlying reason for variation in the selenium content of foods is that the selenium content of soil can vary tremendously; from areas in South Dakota where animals grazing on fields will die within days of selenium toxicity, to areas in New Zealand where animals will die of deficiency unless supplemented. Even within this country, the selenium content of corn can vary by as much as 200-fold due to differences in soil levels (Ullrey, 1981). This, in turn, is reflected in an approximately 50-fold variation in selenium content of swine muscle. Because of the complexities of the food distribution system within most countries, it is extremely difficult to know the ultimate source of any particular slice of bread or cut of meat. Hunter and colleagues (1987) have provided evidence that the calculation of selenium intake is unlikely to provide useful information on intake of this element; calculations of dietary selenium intake (based on a version of the questionnaire used to estimate beta-carotene intake in the example note previously) were not correlated with nail selenium levels even though the use of selenium supplements within the range of normal dietary intakes exhibited a clear dose-response relationship with nail levels.

Beta-carotene and selenium provide clear contrasts of nutrients for which the assumption of constant nutrient content of foods is and is not reasonable; it is not entirely clear where other nutrients lie on this spectrum. For many of the major constituents of foods that are reasonably stable under typical preserva-

tion, storage, and cooking conditions, such as the primary dietary fats, carbohydrate fractions, and calcium, the assumption of constant nutrient content is probably not seriously violated. For some nutrients, such as folic acid which is rather labile, however, it is not yet clear whether calculated intake will realistically reflect true intake. Recognizing that all specimens or examples of any particular food do not necessarily have the same nutrient content, the designers of major food composition databases have attempted to provide increasingly specific information. For example, values are provided for specific cuts of meat, specific methods of food preparation, and specific manufacturers of prepared food. In some cases, this provides improved calculations of nutrient intake; if detailed food information is available, this would usually require data on a meal-by-meal basis. Unfortunately, a high degree of specificity is generally not useful in epidemiologic applications as such detailed information is usually not obtainable for food intake over extended periods of time. In the context of validation studies (see Chapter 6), however, the maximal degree of specificity is generally desirable.

Uncertainty regarding the constancy of the nutrient content of food raises the issue of how this assumption can be evaluated. One classic method of evaluating assumptions used in the calculation of nutrient intake is to collect replicate food samples for single meals or 24-hour periods, meaning that for each food eaten an identical serving is placed in a container for analysis. Nutrient intake is then estimated by calculation from the foods that were recorded as eaten and also by chemical analysis of a homogenate of the duplicate meals consumed by each subject. For example, Moser and Allen (1984) compared zinc analyses of three 24-hour duplicate meals with calculated intakes among 36 lactating or nonlactating postpartum women; calculated and analyzed mean values were nearly identical. Although this method can be very useful for identifying serious problems, the correlations between calculated and chemically analyzed intakes are likely to be overly optimistic as the differences between subjects based on single 24-hour periods will be much larger than the true long-term differences between subjects (see Chapter 3). Some of the differences in nutrient intake between subjects are due to differences in total food intake, thus adjustment for total caloric intake reduces between-subject variation and, therefore, correlation coefficients, even further. For example, it is possible that identical mean values and positive correlations would be seen for selenium using this method due to large intakes of meat, dairy products, or overall food intake on some days, even though other evidence indicates that calculated intakes are unlikely to be useful in realistic epidemiologic settings.

A second and more rigorous method of evaluating the assumption of constant nutrient composition of foods is to examine the correlation of calculated nutrient intakes with a biochemical indicator of nutrient intake, such as for beta-carotene as discussed in the previous example. For a calculated intake and biochemical indicator to be correlated requires that intake of foods be accurately reported, the nutrient content of the foods be known accurately and be constant within specific foods, the nutrients be similarly bioavailable from the different foods, and the level of the biochemical indicator be responsive to nutrient intake over the range being studied. Thus, the demonstration of a positive correlation

provides evidence that the nutrient composition is reasonably accurate and constant within foods. Failure to observe a correlation, however, does not imply that composition data are faulty as any of the other assumptions may not be met. As discussed later, responsive biochemical indicators of intake do not exist for most nutrients of major interest, thus seriously limiting the application of this approach.

Given the lack of more direct information, judgments about the appropriateness of calculated intakes often needs to be made indirectly based on knowledge about the stability of nutrients and their variability in foods. For example, folic acid tends to be less stable during preservation, processing, and cooking than many nutrients, thus raising uncertainty about the validity of calculated values. For such a nutrient, the findings of epidemiologic studies need to be interpreted cautiously as absence of an association with disease could be due to large variation in the nutrient content of specific foods.

The calculation of nutrient intakes from information on food consumption requires two resources: the food composition data itself and computer software to perform the calculations (hand calculation now being obsolete). These are discussed separately.

Food composition information is needed for two general purposes in epidemiologic studies. The first is for the analysis of traditional open-ended dietary data collected by short-term recall or meal-by-meal recording methods (see Chapter 4). This requires an extensive and comprehensive database as nutrient values are needed for all foods that might be reported by subjects. In contrast, most epidemiologic studies employ a structured questionnaire consisting of not more than 100 or 200 food items (see Chapter 5). To compute nutrient intakes based on information obtained with such a questionnaire, a customized nutrient database must be created to provide values for each nutrient to be computed for every food on the questionnaire. Investigators need to compile this information, often by using an available comprehensive database supplemented with other sources of information as needed. Whether an existing comprehensive or a custom database is used, several features need to be considered:

1. Most fundamentally, food composition data should be as accurate and as up-to-date as possible. Not only are technologic advances making improved analyses possible, but also food composition itself changes over time due to selective plant and animal breeding and alterations in preservation and processing techniques. New foods and formulations are constantly being added to the market.
2. Uniformity in the source of nutrient composition is desirable as it is important that, for any one nutrient, all foods are analyzed by the same method.
3. Comprehensiveness in the scope of foods is particularly important when the database is used to analyze open-ended data on food intake, as all foods reported must be assigned nutrient values.
4. Specificity may be especially important for some nutrients that are affected by manufacturing or processing. For example, if sodium intake is of particular interest when analyzing open-ended food intake informa-

tion, considerable specificity may be needed regarding brands and types of processed food.

5. Completeness in nutrient values is of extreme importance when a database is used to compute nutrient intakes; no food should have a blank value as this will effectively be counted as zero when summarizing intake of that nutrient for a person. This issue represents a continuing challenge for persons maintaining a database because, even for a single nutrient, food analysis laboratories rarely conduct measurements of all foods that are listed in any single database. For example, cholesterol measurements may well not be made on both fried eggs and poached eggs, although these may be listed as separate items in the database. Some database systems have not been willing to make the reasonable assumption that cholesterol content is the same in both these items, even though only one has been analyzed directly. Although blank values may be perfectly appropriate for listings of food composition that are purely for reference purposes, this is not tolerable in the case of computerized databases that are used to calculate intakes as an intelligently imputed value based on analyses of similar foods is almost always a better approximation to truth than a zero value.
6. Comprehensiveness in the nutrients for which values are provided is a highly desirable feature. No database can be truly complete because technology continually provides additional ways to subdivide and characterize the components of our food supply. Because it is rarely appropriate to examine only a few nutrients in isolation in epidemiologic studies, a database should provide the ability to calculate simultaneously at least the major components of our food.

Specific Sources of Dietary Data

In constructing a food composition database for a specific application, it is typically necessary to use multiple sources of information. For example, to compile values for the questionnaire used in a prospective study among U.S. nurses (Willett et al., 1987), nutrient information was obtained from several governmental and commercial sources, approximately 15 different manufacturers and 27 journal articles; this list continues to grow as the scope of nutrients computed increases. Available sources of food composition data include the following:

1. *United States Department of Agriculture (USDA) publications.* The USDA provides comprehensive and up-to-date information on food composition. This resource is almost always the starting point for any specific database; an outline of available documents is included in Table 2–1. The central publication is *Composition of Foods—Raw, Processed, Prepared,* Agriculture Handbook No. 8 (U.S. Dept. of Agriculture, 1963). This is published as a series of revised sections with updated values for one or two food groups released annually. This information is also available on magnetic tape (e.g., Nutrient Data Base for Standard Reference, Release 6, 1987; available from the National Technical Information Service, U.S.

Table 2–1. Sources of food composition data

USDA Publications[a]

1. Composition of Foods—raw, processed, prepared
   Agricultural Handbook No. 8, 1963

   Revised Sections of Agricultural Handbook No. 8:
   - 8-1 Dairy and egg products. 1976.
   - 8-2 Spices and herbs. 1977.
   - 8-3 Baby foods. 1978.
   - 8-4 Fats and oils. 1979.
   - 8-5 Poultry products. 1979.
   - 8-6 Soups, sauces and gravies. 1980.
   - 8-7 Sausages and luncheon meats. 1980.
   - 8-8 Breakfast cereals. 1982.
   - 8-9 Fruits and fruit juices. 1982.
   - 8-10 Pork products. 1983.
   - 8-11 Vegetables and vegetable products. 1984.
   - 8-12 Nut and seed products. 1984.
   - 8-13 Beef products. 1986.
   - 8-14 Beverages. 1986.

2. Nutritive Value of American Foods in Common Units
   Agriculture Handbook No. 456, 1975.

3. Nutritive Value of Foods
   Home and Garden Bulletin No. 72, revised, 1985.

4. Iron Content of Food
   Home Economics Research Report No. 45, 1983.

5. The Sodium Content of Your Food
   Home and Garden Bulletin No. 233, 1980.

6. Provisional Tables:

   Nutrient Content of Canned and Frozen Vegetables, 1979
   Supplement on Folacin Content of Foods, 1979
   Nutrient Content of Bakery Foods and Related Items, 1981
   Nutrient Content of Canned, Dried, and Frozen Fruit, 1981
   Nutrient Content of Beverages, 1982
   Amino Acids in Fruits and Vegetables, 1983
   Fatty acid and Cholesterol Content of Selected Foods, 1984
   Nutrient Content of Fast Foods, 1984
   Percent Retention of Nutrients in Food Preparation, 1984
   Omega-3 Fatty Acids and Other Fat Components in Selected Foods, 1986
   Sugar Content of Selected Foods, 1986
   Vitamin K Content of Selected Foods, 1987

Non-USDA Food Composition Tables

McCance and Widdowson's The Composition of Foods
Paul and Southgate, 1978

Bowes and Church's Food Values of Portions Commonly Used
Pennington and Church, 1985

Nutritive Value of Convenience and Processed Foods
West Suburban District of the Illinois Dietetic Association and the American Dietetic Association; Chicago, IL, 1987

[a]Available from Superintendent of Documents, U.S. Govt. Printing Office Washington, D.C., 20402.

Department of Commerce, Springfield, VA 22161). Databases used for the National Food Consumption Survey and the National Health and Nutrition Surveys include data for a large number of recipes and are also available on tape from the National Technical Information Service.

2. *Non-USDA food composition tables.* A variety of food composition databases have been published by other governmental units, or commercial establishments. Several of the more commonly used references are noted in Table 2–1.
3. *Manufacturer's data.* In some instances information obtained from food manufacturers is useful. Data on specific brands of breakfast cereals may be particularly important because of extensive fortification with vitamins and minerals and because of their regular use by many persons. Names and addresses of food manufacturers are published periodically (Thomas Publishing Co., 1985a, b).
4. *Professional journals.* For some nutrients that have not yet been incorporated into the major databases, it is sometimes necessary to use primary published sources of food analyses.

Nutrient Computation Systems

The computation of nutrient intakes from food consumption information requires both a computerized food composition database and the software to perform these calculations. The coding and entry of open-ended food information, such as that obtained from a short-term recall, has been an extremely tedious and expensive process requiring a highly trained person, preferably a dietitian. Typically, every food is looked up in a book or list to obtain a code number, which is then entered into a computer. A unit, such an ounce or cup, is then designated and a multiplier is entered. Such time-consuming and error-prone procedures have rapidly become archaic with the availability of inexpensive computing resources. Recently several automated coding systems have been developed that allow direct entry of foods by name, avoiding the manual entering of numerical codes by the user.

The available choices in food composition databases and nutrient analysis systems are presently in a state of rapid evolution, therefore, details of these are not provided. A directory of nutrient data banks has been assembled by Hoover (1987), which is updated annually. Features of several of the larger and more commonly used systems that include both a database and some form of analysis software are included in Table 2–2. The features of available analysis systems are also summarized in the *Annual Journal of Dietetic Software* (this may be purchased from JDS, P.O. Box 2565 Norman, OK 73070).

In choosing a nutrient computation system, all of the points noted previously regarding the database employed should be considered. In addition, several factors related to the software should be considered:

1. The manual entry of coding numbers for foods should be avoided if possible; as discussed, coding should be provided as an automated feature using an interactive computer system.

Table 2–2. Examples of nutrient computation systems

Systems	No. of foods	Nutrients per food	Portion conversion capabilities	Food entry	Services	Computer
Case Western Reserve University	3000	71	Many	Manual numerical coding	Sell, rent processing	Vax AT&T
Ohio State University	9100	68	Many	Manual numerical coding	Sell, rent processing	IBM
Nutrition Coordinating Center (Minnesota)	1800	71	Many	Manual numerical coding	Sell, processing	IBM
University of Massachusetts	6000	53	Many	Manual numerical coding	Rent, processing	CDC
CBORD[a]	4000	up to 100	Many	Direct entry of foods	Sell, rent processing	IBM-PC

[a]Provides software to use other databases (e.g., Ohio State University Database).
From Hoover, L. W., 1987.

2. The system should have the capacity to add foods or nutrients of particular interest for a specific study, and to update or modify existing values.
3. The system should provide flexibility for different portion units so that the user does not have to convert manually various measures of volume or weight.
4. The system should provide an option to display nutrients calculated for each food, as well as summaries for each meal or day.
5. The system should allow common recipes to be compiled and entered as a single item, rather than reentering each of the components of the recipe separately.
6. Transfer of data to other computers where the information will ultimately be used should be easily accomodated.

At present, many of the most commonly used comprehensive nutrient analysis systems reside on mainframe computers. Some, like the Minnesota Nutrition Coordinating Center, primarily process information that is submitted to the central unit, which allows strict standardization of coding procedures. The availability of inexpensive hard disks on personal computers, however, now makes it possible to maintain not only the entry systems, but also a comprehensive database and analysis systems on such machines. Although efficient systems for the analysis of open-ended dietary data will become increasingly available, this will not reduce the need for careful training and monitoring of the personnel who use them.

REFERENCES

Ames, B. N., R. Magaw, and L. S. Gold (1987). Ranking possible carcinogenic hazards. *Science 236*, 271–280.

Busby, W. F., Jr. and G. N. Wogan (1984). Aflatoxins. In Searle, C. E., ed., *Chemical*

Carcinogens, 2nd ed., Vol. 2. Washington, D.C.: ACS Monograph 182, American Chemical Society, pp. 945–1136.

Doll, R. and R. Peto (1981). The causes of cancer: Quantitative estimates of avoidable risks of cancer in the United States today. *J.N.C.I. 66,* 1191–308.

Graham, S., H. Dayal, J. Swanson, A. Mittelman, and G. Wilkinson (1978). Diet in the epidemiology of cancer of the colon and rectum. *J.N.C.I. 61,* 709–714.

Hoover, L. W., ed. (1987). Nutrient Data Bank Directory, 6th ed., Houston, TX.: Twelfth Annual National Nutrient Data Bank Conference.

Hunter, D. J., C. G. Chute, E. Kushner, G. A. Colditz, M. J. Stampfer, F. E. Speizer, J. S. Morris, and W. C. Willett (1987). Predictors of selenium concentration in nail tissue (abs.). *Am. J. Epidemiol. 126,* 743.

Jacobson, H. N. and J. L. Stanton (1986). Pattern analysis in nutrition. *Clin Nutr 5,* 249–253.

Lands, W.E.M. (1986). *Fish and Human Health.* Orlando: Academic Press.

Manousos, O., N. E. Day, D. Trichopoulos, F. Gerovassilis, A. Tzanou, A. Polychronopoulou (1983). Diet and colorectal cancer: a case-control study in Greece. *Int. J. Cancer 32,* 1–5.

Mertz W. (1981). The essential trace elements. *Science 213,* 1332–1338.

Mertz W. (1984). Foods and nutrients. *J. Am. Diet. Assoc. 84,* 769–770.

Morris, J. N., J. W. Marr, and D. G. Clayton (1977). Diet and heart: A postscript. *Br. Med. J. 2,* 1307–1314.

Moser P. B. and D. Allen (1984). Zinc intakes of lactating and non-lactating women; Analyzed vs. calculated values. *J. Am. Diet. Assoc. 84,* 42–46.

Peto, R., R. Doll, J. D. Buckley, and M. B. Sporn (1981). Can dietary beta-carotene materially reduce human cancer rates? *Nature 290,* 201–208.

Sugimura, T. (1986). Studies on environmental chemical carcinogenesis in Japan. *Science 233,* 312–318.

Thomas Publishing Company, (1985a). *Thomas Grocery Register. Vol. 1. A–Z Index and Trademarks/Brandnames.* New York: Thomas Publishing Co.

Thomas Publishing Company, (1985b). *Thomas Grocery Register. Vol. 2. Products and Services.* New York: Thomas Publishing Co.

Troll, W., K. Frenkel, and R. Wiesner (1984). Protease inhibitors as anticarcinogens. *J.N.C.I. 73,* 1245–1250.

Ullrey, D. E. (1981). Selenium in the soil-plant-food chain. In Spallholz, J. E. et al., eds., *Selenium in Biology and Medicine.* Westpoint, Conn.: AVI Publishing; pp. 176–191.

U.S. Dept. of Agriculture, (1963). Composition of foods—raw, processed, and prepared. Agricultural Handbook No. 8. Washington, D.C.: U.S. Gov't. Printing Offices.

Wattenberg, L. W. and W. D. Loub (1978). Inhibition of polycyclic aromatic hydrocarbon-induced neoplasia by naturally occuring indoles. *Cancer Res. 38,* 1410–1413.

Welsh, S. O. and R. M. Marston (1982). Review on trends in food use in the United States, 1909 to 1980. *J. Am. Diet. Assoc. 81,* 120–128.

Willett, W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, F. E. Speizer (1987). Dietary fat and risk of breast cancer. *N. Engl. J. Med. 316,* 22–28.

Willett, W., L. Sampson, M. L. Browne, M. J. Stampfer, B. Rosner, C. H. Hennekens, and F. E. Speizer (1988). The use of self-administered questionnaire to assess diet four years in the past. *Am. J. Epidemiol. 127,* 188–199.

Witteman, J.C.M., W. C. Willett, M. J. Stampfer, G. Colditz, F. Sacks, B. Rosner, F. Speizer, and C. Hennekens (1987). Dietary calcium and magnesium and hypertension: A prospective study (abstr.). *Circulation Supplement 76: (s-4),* iv–35.

3

Nature of Variation in Diet

For most epidemiologic applications, long-term diet, rather than intake on any specific day or small number of days, is the conceptually relevant exposure. The period of time may be years in studies of factors that affect atherogenesis or cancer, or a critical period of a few weeks in studies of nutrients that influence fetal malformation. In studies of physiologic intermediates, such as plasma lipids or endogenous hormone excretion, dietary intake during several days, weeks, or months may be of interest. Because we are usually interested in long-term diet, an understanding of day-to-day variation in dietary intake is essential in choosing an appropriate method to assess diet and to interpret data collected by various approaches.

A central feature of the dietary intake of free-living individuals is variation from day to day superimposed on an underlying consistent pattern. (If there were no element of consistency, and daily intake were a completely random event, there would be no hope of measuring the effects of nutrients epidemiologically.) A number of factors, such as day of the week or season, may contribute to daily variation in dietary intake in a systematic manner. The magnitude of these influences is largely determined by cultural and ecologic factors. For example, it is traditional for many American families to have unusually large meals on Sunday. In countries without extensive food preservation and transportation systems, seasonal effects are relatively strong (Brown et al., 1982); for example, in some parts of the world much of the vitamin A is consumed during a limited portion of the year when certain fruits and vegetables are available. In industrialized societies, seasons make a relatively small contribution to variation in nutrient intake (van Staveren et al., 1986b); however, intakes of some fruits and vegetables may vary substantially according to the time of year (Ziegler et al., 1987). Although there is evidence that total caloric intake varies with the menstrual cycle (Dalvit, 1981), most of the variation in an individual's diet is without an obvious pattern. This apparently random variation is largely due to true variation in the food that is eaten, but also has a component of measurement

error, meaning error in the measurement of food intake on a given day. Because the topic of measurement error is discussed in detail in other chapters, this source of variation is ignored for the moment.

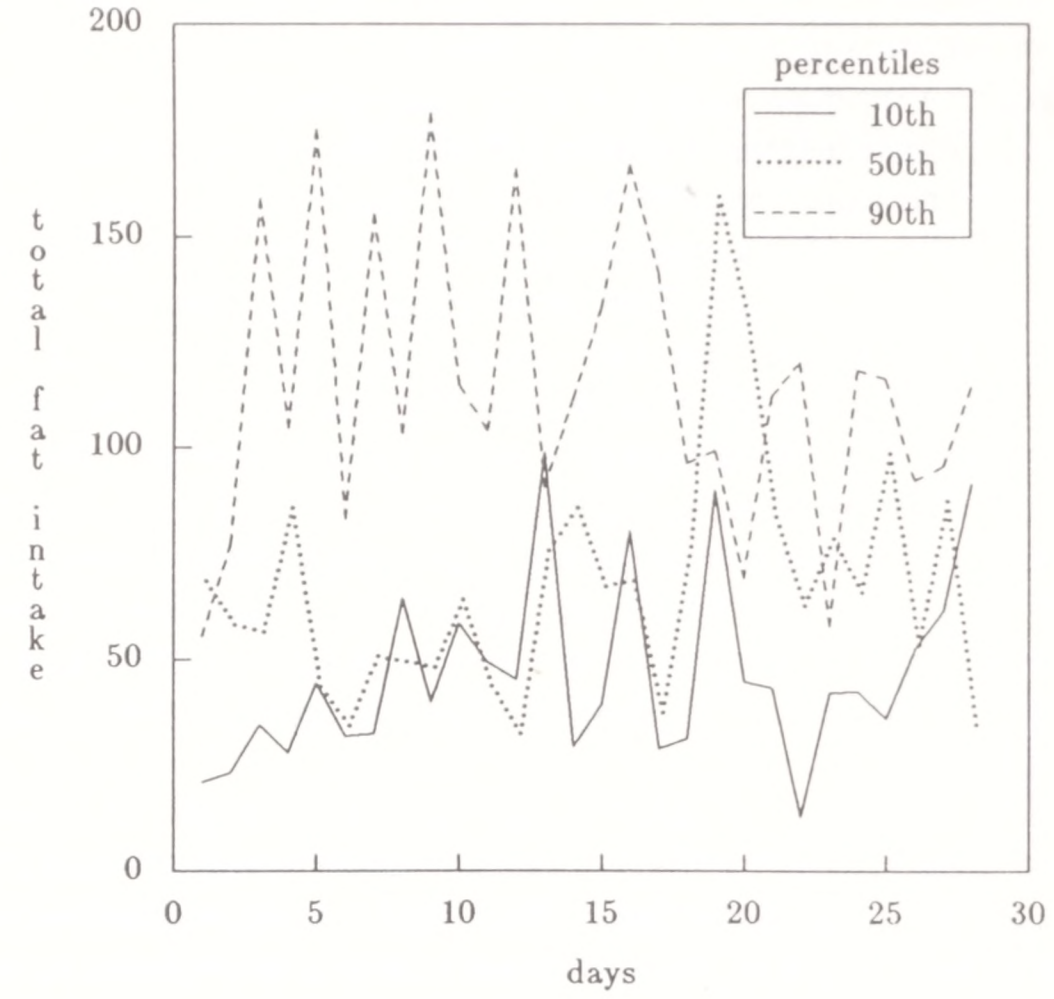
The degree of random variation differs according to nutrient. For macronutrients, which make a large contribution to total caloric intake and thus may have fairly constrained physiologic limits, there is less possibility for large degrees of variation. As micronutrients tend to be concentrated in certain foods, intake can be very low or very high, depending on food choices for that day. The daily variation in total fat and total dietary vitamin A intake for three women (one from each of the high, middle, and low quintiles of the 28-day average intakes of these nutrients) is displayed in Figure 3–1. The data for these and the other examples are based on diaries provided by 194 Boston-area women who recorded their daily food intake during four 1-week periods over a year (Willett et al., 1985). It can be appreciated that on any single day, the individual intakes overlap considerably, even though on the average they are distinct. It can be seen that variability is greater for vitamin A than for total fat, and is also related to the level of intake.

The consumption of a specific nutrient for a single individual can be described as a frequency distribution of daily intakes with a mean and standard deviation (Fig. 3–2); the individual's *true* intake can be considered as the mean for a large number of days. In most epidemiologic investigations we are interested in measuring these true intakes for individual subjects. Collectively, these true individual intakes define a frequency distribution for the study population as a whole; this distribution characterizes the true population mean and standard deviation. Distributions for mean daily intakes of total calories, total fat, and vitamin A for the group of 194 women are shown in Figure 3–3. This figure closely describes the true between-person variation for this population as the day-to-day fluctuations will be dampened or canceled out by averaging many days for each person, in this example 28 days.

In reality, it is rarely possible to measure a large number of days of dietary intake for an individual subject; therefore, intakes during a sample of one or several days are usually measured. The effect of this sampling on the apparent distribution of intakes for individual subjects will be to artificially increase the standard deviation (i.e., to broaden the tails of the distribution). This effect can be appreciated by considering that individuals with the highest true intakes have some days when their intake is higher than their long-term average, and individuals with the lowest true intake have days when their intake is below their long-term mean. Thus, the observed distribution has exteme values that are higher and lower than any of the true long-term averages for any subject. To graphically demonstrate the effect of sampling a small number of days for an individual on the apparent between-person distribution of dietary intake, data from the 194 Boston-area women who each recorded daily food intake during four 1-week periods are shown in Figure 3–4.

The effect is fairly dramatic. In data based on a single day per subject, those at the 90th percentile were consuming 3.0 times as much fat and 6.4 times as much total vitamin A (excluding supplements) as were those at the tenth percentile. Using data based on 1 week of intake, these ratios were 2.2 for fat and

A



B

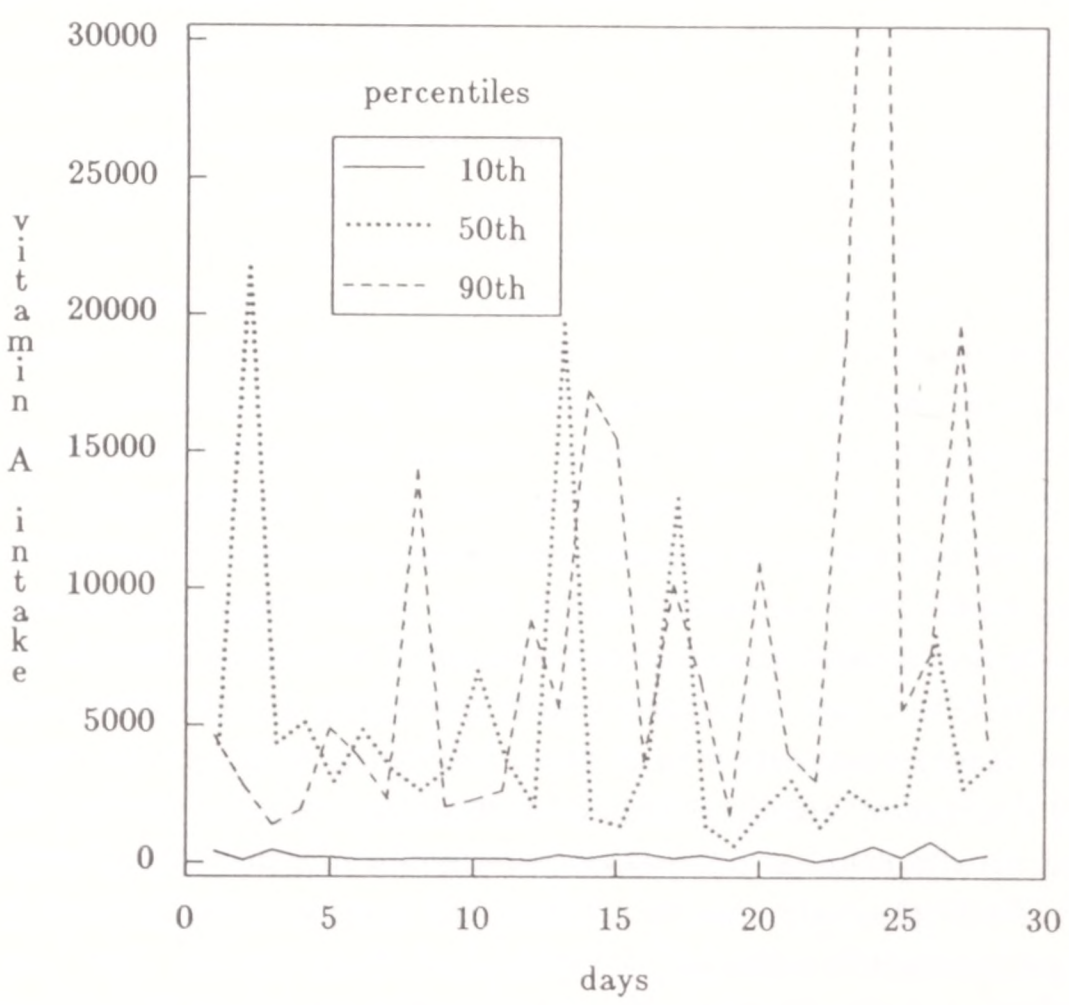


Figure 3–1. Daily intakes for three women at the 10th, 50th, and 90th percentile of distributions for total fat intake in grams (**A**), and vitamin A intake in international units (**B**). See text for details.

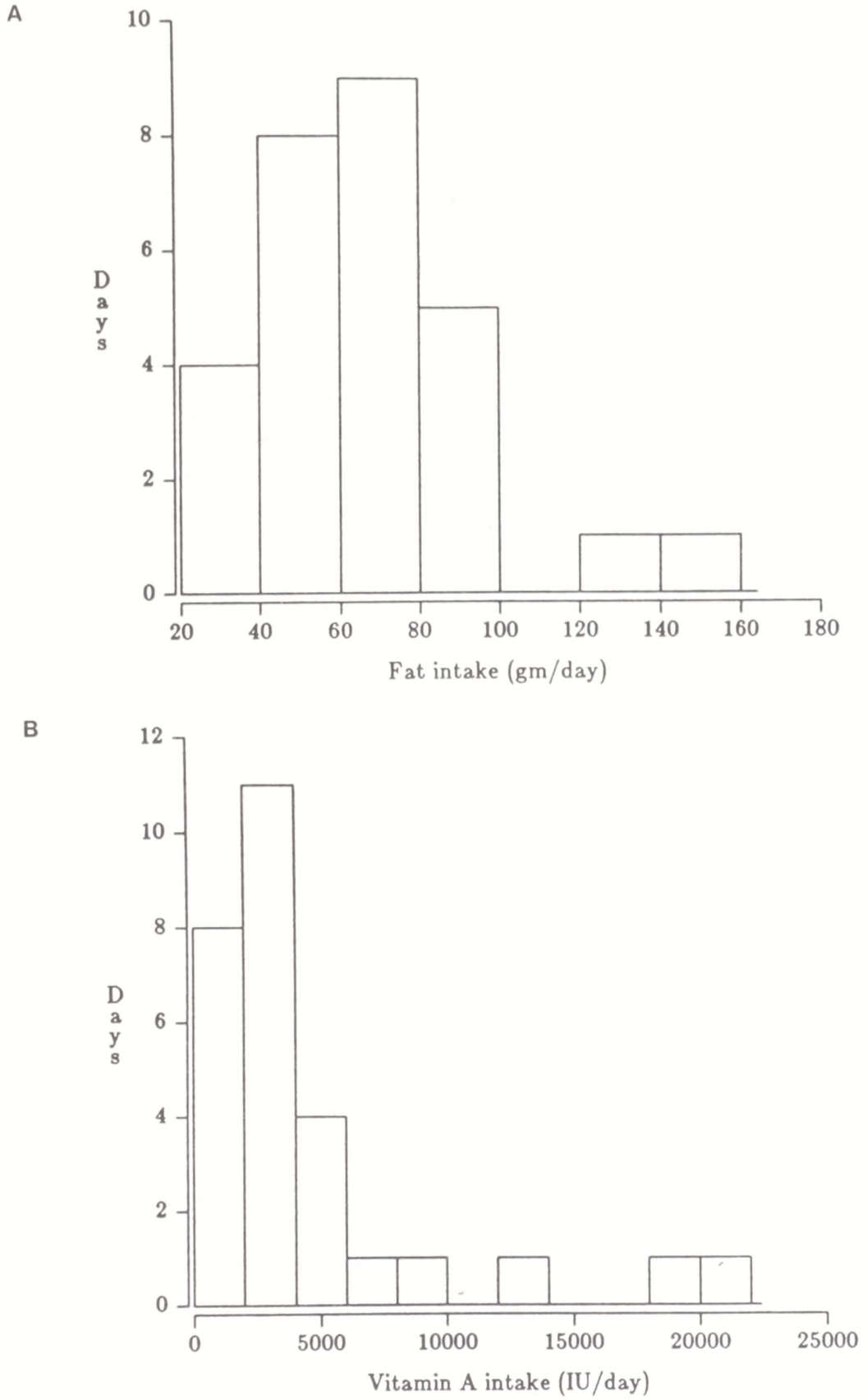
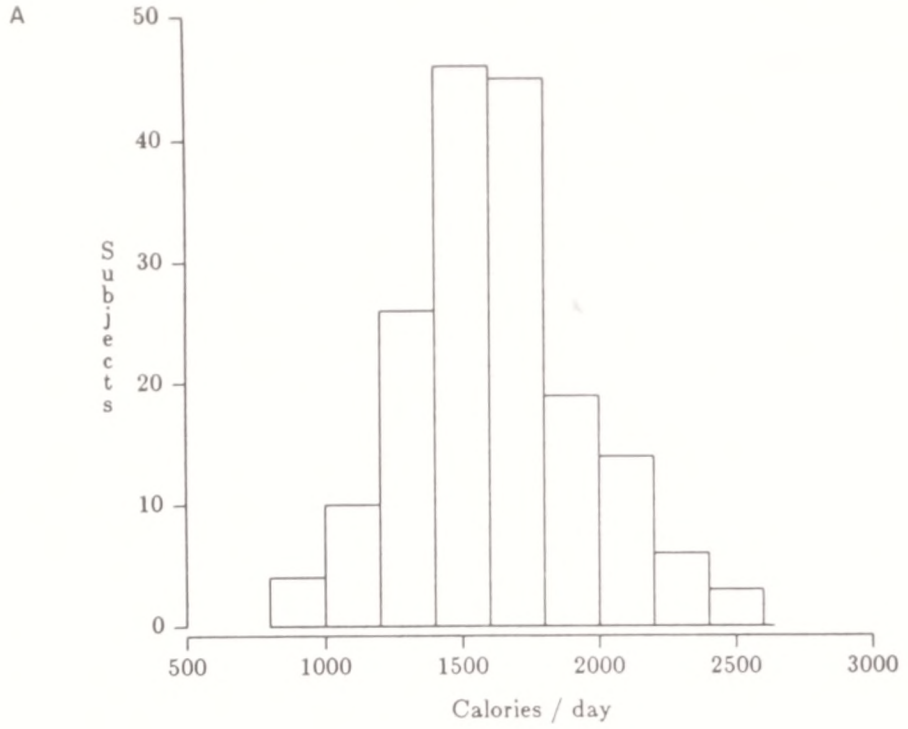
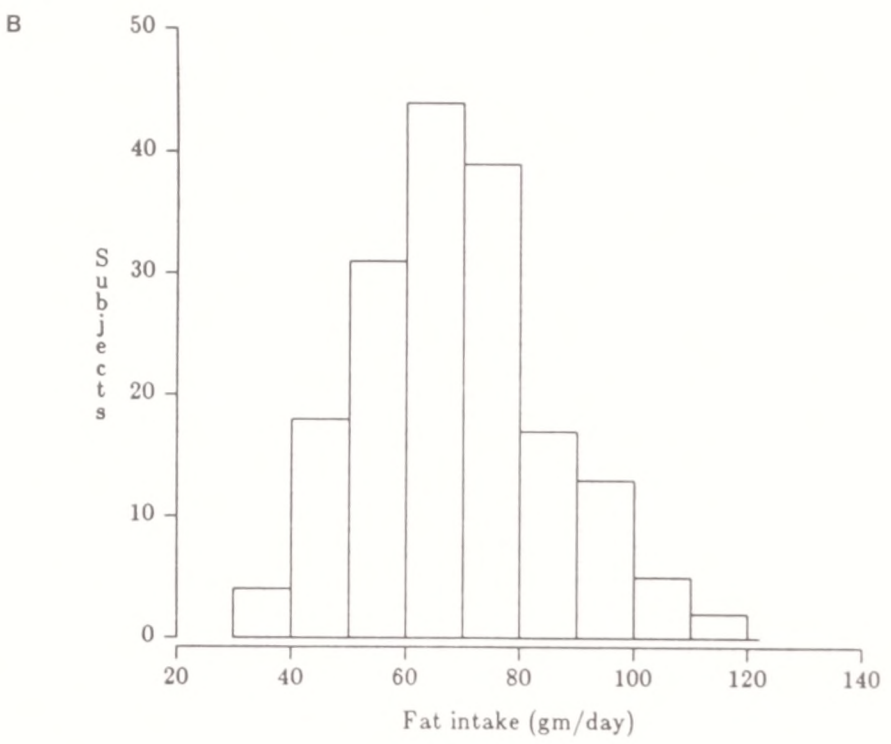


Figure 3–2. Frequency distribution of daily intakes of total fat (**A**), and vitamin A (**B**) for one woman (the women at the 50th percentile in Fig. 3–1). For this person, 28 days of intake are displayed; the mean of these days can be considered as the true intake for that person.

3.0 for vitamin A. The change observed using a 4-week sample was less striking; the ratios were 1.9 for fat and 2.5 for vitamin A, suggesting that a further increase in the number of days per subject would have a minimal effect on the distribution. The distortion of true between-person variation in dietary intake resulting from the use of a single day of dietary intake per subject is not purely





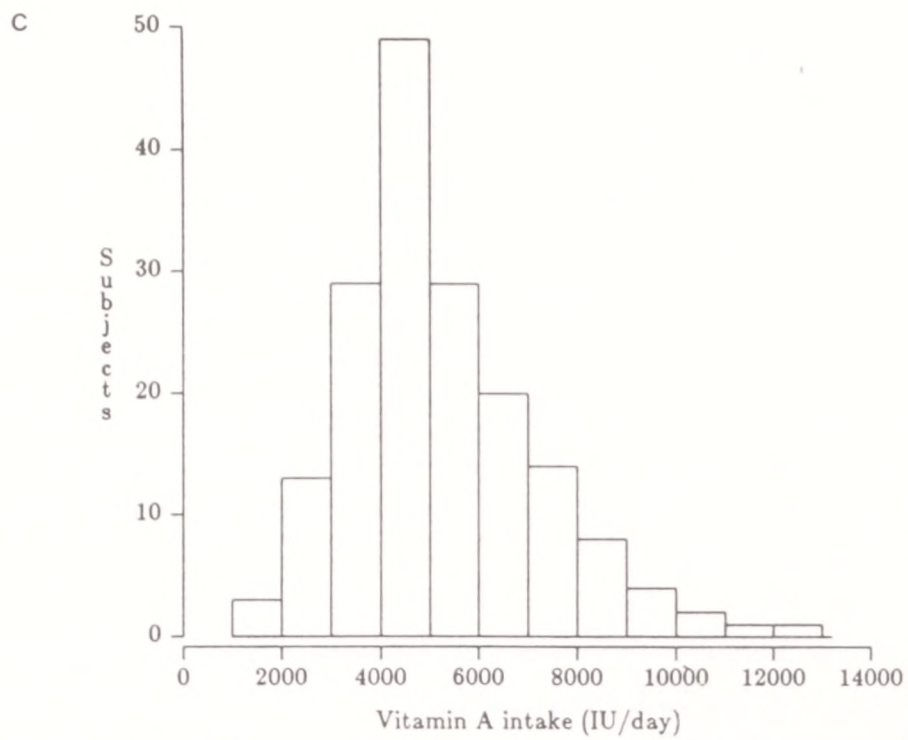


Figure 3–3. Distributions of average daily intakes (based on 28 days per person) for calories (**A**), total fat (**B**), and vitamin A (**C**) for 194 women. Because the average of a large number of days is used for each woman, these distributions closely approximate the true between-person variation for this group of women.

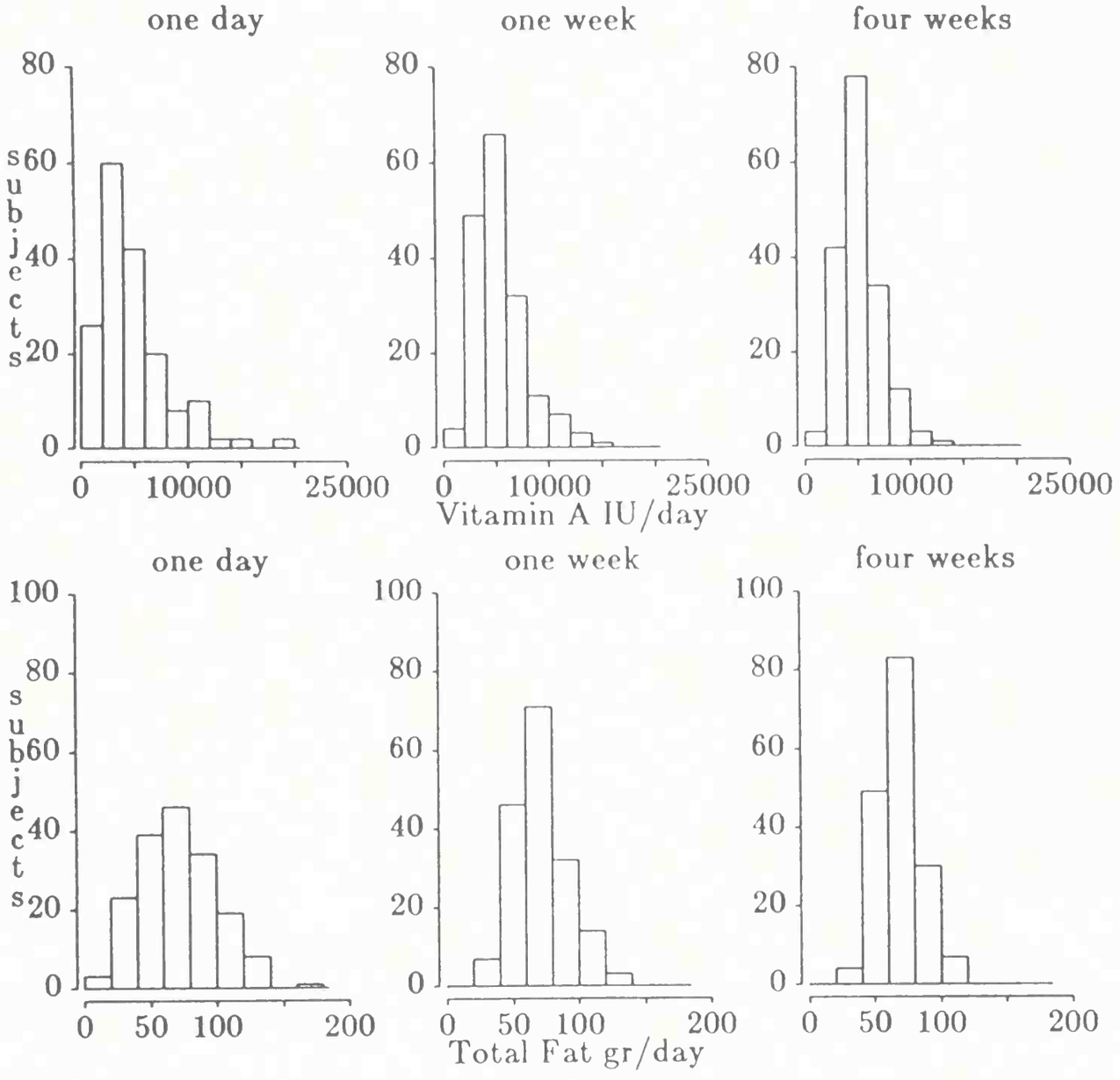


Figure 3–4. Effect of sampling 1 day, 7 days, or 28 days per subject on the observed distribution of fat and vitamin A intake. Data are based on diet records provided by 194 women.

an academic matter. National nutrition surveys have traditionally employed only a single 24-hour recall per subject; thus many major reports of the distributions of dietary intake are extremely misleading. For example, it is not really plausible that 5 percent of U.S. men aged 25–34 years consume less than 1250 calories per day and that 10 percent consume more than 4000 calories per day, as suggested by a publication based on the first National Health and Nutrition Survey (NHANES I) (U.S. Dept HEW, 1979). Distributions for micronutrients tend to be even more seriously distorted due to their greater day-to-day variation.

The variation in daily intake of specific nutrients has been studied formally by Beaton and colleagues (1979 and 1983), Liu and colleagues (1978), Rush and Kristal (1982), Sempos and colleagues (1985), and Hunt and colleagues (1983) using analysis of variance techniques. In this approach, the daily nutrient intake is the dependent (outcome) variable and explanatory variables are the independent variables. Because within-person intake is being examined, at least 2 days of information are needed for each person. The data are arrayed as shown in

Table 3–1. Data arrayed for analysis of variance (ANOVA) using the model described in Equation (3-1)

Nutrient (Y)	Person	Factor X	Day of week
value 1	1	1	1
value 2	1	1	2
value 3	1	2	1
value 4	1	2	2
value 5	2	1	1
value 6	2	1	2
value 7	2	2	1
value 8	2	2	2
value 9	3	1	1
value 10	3	1	2

Table 3–1 with repeated measurements for persons considered as separate "records."

Data are analyzed using a random effects (repeated measures) model of the form:

$$\text{Nutrient } Y_{ijk} = \mu + \text{subject}_i + \text{factor } X_{ij} + \text{day of week}_{ijk} + \epsilon_{ijk}. \quad (3\text{-}1)$$

In this model, the long-term average intake can vary between subjects, but within a subject daily diet can be influenced by specific identifiable variables denoted by factor X_{ij} such as season and day of the week. The error term (ϵ_{ijk}) represents the random within-person variance, that is, the day-to-day variance within a person not explained by the other independent variables.

Among a group of adult men and women, Beaton and colleagues (1979) found that the within-person and between-person factors were the major contributors to variance for all nutrients examined (Tables 3–2 and 3–3). Days of

Table 3–2. Relative sources of variation for daily nutrient intake, not adjusted for energy intake

Component of variance	Energy	Protein	Carbohydrate	Fat	SFA	MFA	PFA	Cholesterol
Men								
Subjects	51.8%	39.2%	42.0%	48.6%	52.0%	42.8%	27.0%	25.7%
Sequence	0.4%	0	0	1.2%	0.4%	1.3%	1.1%	0.6%
Interviewer	0	0	0.5%	0	0	0	0	0
Day of week	0	0	0	0	0	0	0	0.6%
Residual	47.7%	60.8%	57.5%	50.2%	47.6%	55.9%	72.0%	73.2%
Women								
Subjects	41.9%	38.9%	44.3%	38.3%	41.3%	37.4%	21.2%	19.1%
Sequence	0	0	0	0.3%	0	1.4%	1.3%	0
Interviewer	0	0	0	0	0	0	0	0.6%
Day of week	9.4%	5.3%	4.5%	5.7%	4.2%	5.2%	3.2%	4.6%
Residual	49.4%	55.7%	51.2%	55.7%	54.6%	56.1%	74.3%	75.8%

SFA, saturated fatty acids; MFA, monounsaturated fatty acids; PFA, polyunsaturated fatty acids.
From Beaton et al., 1979, reproduced with permission.

Table 3–3. Relative sources of variation for daily nutrient intake divided by energy intake

Component of variance	Protein	Carbohydrate	Fat	SFA	MFA	PFA	P:S ratio	Cholesterol mg/1000 kcal	Cholesterol mg/1000 g fat
Men									
Subjects	13.7%	31.3%	20.8%	23.2%	19.8%	16.4%	20.2%	10.5%	6.2%
Sequence	0	0	0.5%	0	0	1.7%	0.6%	0	0
Interviewer	0.3%	0	0	0	0	0.8%	0	0	0
Day of week	0	0	0	0	0	0	0	0	0
Residual	86.0%	68.7%	78.7%	76.8%	80.2%	81.1%	79.2%	89.5%	93.8%
Women									
Subjects	18.5%	37.2%	30.0%	33.9%	24.7%	8.8%	11.3%	12.8%	18.3%
Sequence	0	0	0	0	0	2.1%	1.7%	0.8%	0.8%
Interviewer	0	0.7%	1.3%	0	2.2%	0	0	1.4%	3.3%
Day of week	0	0	0	0	0	0.4%	0	0	0
Residual	81.5%	62.1%	68.7%	66.1%	73.1%	88.8%	87.0%	85.0%	77.6%

SFA, saturated fatty acids; MFA, monounsaturated fatty acids; PFA, polyunsaturated fatty acids.
From Beaton et al., 1979, reproduced with permission.

the week (with intakes higher for Sundays) explained a small portion of variance for women but not for men. Different interviewers and sequence of days of data collection, however, made negligible contributions. Thus, dietary intake could be reasonably well expressed by the simple model where ϵ represents the day-to-day variation:

$$\text{Nutrient } Y = \mu + \text{subject}_i + \epsilon \tag{3-2}$$

For total caloric intake, the within-person variance was approximately equal to the between-person variance, but for specific nutrients the ratio of within-person to between-person variance was greater than one; for polyunsaturated fat, cholesterol, and most micronutrients the within-person variance was much greater. When nutrient intakes were considered in relation to total energy intake, expressed as a percentage of total calories, the differences between persons decreased so that the relative importance of the within-person component of variance was even larger, ranging from approximately 30 to 90 percent of total variance for the nutrients examined.

As pointed out by Beaton and co-workers, the pattern of high within-person variation in nutrient intakes is largely cultural and these findings may not necessarily apply to other populations. In other U.S. studies, however, similarly high within-person variation has been observed (Table 3–4).

The contribution of variable intervals of time between repeated measurements to the variation in dietary intake deserves further consideration. In one unpublished analysis among postmenopausal women, we observed less within-person variation when the days were consecutive than when separated by several months. This could have important implications for study design as misleading estimates of within-person variation may be obtained using consec-

Table 3–4. Ratios of the within-person and between-person components of variance (S_w^2/S_b^2) for nutrient intakes observed in several North American studies

	Beaton et al. (1983)		Liu et al. (1978)		Hunt et al. (1983)		Sempos et al. (1985)	NHS[c]	Rush et al. (1982)
Nutrient	*20 Men*	*30 Women*	*181 Men*[a]	*318 Men*[b]	*25 Men*	*25 Women*	*151 Women*	*173 Women*	*225 pregnant women*
Energy (calories)	1.0	1.4	1.8	2.2	1.0	0.8	1.6	1.9	1.1
Protein	1.4	1.4	—	—	1.2	1.3	2.1	3.9	1.4
(% of calories)	5.8	4.0	—	—	2.3	1.9	—	2.7	—
Carbohydrate	1.7	1.4	—	—	2.0	1.2	—	1.2	1.2
(% of calories)	2.3	1.7	—	—	1.9	1.6	—	1.9	—
Total fat	1.2	1.7	—	—	1.2	0.9	—	2.8	1.2
(% of calories)	4.8	2.6	2.3	1.3	1.5	1.5	—	4.1	—
Saturated fat	1.0	1.4	—	—	2.2	1.7	—	2.8	—
(% of calories)	3.2	2.0	2.6	3.9	—	—	—	4.1	—
Polyunsaturated fat	2.9	4.0	—	—	3.5	2.2	5.0	5.0	—
(% of calories)	5.3	7.8	—	—	—	—	6.2	6.2	—
Cholesterol	3.6	4.4	3.8	1.8	5.6	4.2	6.8	6.8	—
per 1000 calories	3.6	2.7	—	—	—	—	5.7	5.7	—
Vitamin C	4.0	2.3	—	—	2.3	2.8	2.3	2.0	—
Vitamin B_6	—	—	—	—	2.1	3.1	1.9	2.7	—
Folic acid	—	—	—	—	1.6	2.0	1.7	6.5	—
Vitamin A	>100	47.6	—	—	1.6	2.5	3.8	11.7	—
Iron	3.6	2.6	—	—	1.8	1.5	2.7	3.1	—
Calcium	2.6	2.3	—	—	1.1	1.7	1.1	2.2	1.0
Potassium	—	—	—	—	0.9	1.2	1.9	1.9	—
Zinc	—	—	—	—	2.7	1.7	2.2	11.7	—

[a]Japanese men in Japan.

[b]Japanese men in Hawaii.

[c]Unpublished data from analyses of diet records collected as part of the Nurses Health Study questionnaire validation study (Willett et al., 1985).

utively collected diet recalls or records. In addition, the seasonal contribution to variance has not been explored carefully. As mentioned previously, seasonal variation in diet may not be large within the United States and other industrialized countries. In one study, the correlations between 1-week diet records did not vary appreciably when they were collected at intervals of 3, 6, 9, or 12 months (Willett et al., 1985). Year-to-year variation was examined by Hunt and colleagues (1983); for most nutrients the contribution was minor at a 1-year interval in comparison with day-to-day variation.

A basic assumption underlying the analytic methods described previously is that the within-person component of variation is random, that is, for one person, the deviation from their long-term average intake on 1 day is independent from their deviation on the previous day. This assumption has been questioned by El Lozy (1983), who pointed out that humans are subject to homeostatic mechanisms (probably both physiologic and cultural) such that overeating on 1 day is likely to be followed by undereating the next. Morgan and colleagues (1987) formally explored the assumption of independence for consecutive days of dietary intake of energy, fat, vitamin A, and iron among 100 women. They found auto-correlations for many subjects, meaning that intake on 1 day added to the prediction of intake on the next, above and beyond the mean intake for a person. A simple pattern, however, was not evident as persons with positive auto-correlations, indicating high intake on 1 day was associated with high intake on the next, were similar in number to those with negative auto-correlations, indicating that high intake on 1 day was associated with lower intake on the next. Morgan and co-workers also observed, as suggested in Figure 3–1B, that those with higher mean intakes have greater within-person variation. For this reason, it may be useful to transform data, such as by taking the natural logarithm, before proceeding with analyses. Further investigation of the influence of 1 day's diet on intake on the following day may eventually refine our understanding of dietary variation, but is not likely to affect substantially the implications of data already published. The possible nonindependence of intake on consecutive days, however, argues for sampling days at random intervals whenever possible.

Number of Days Necessary to Estimate True Intake

A single day provides a poor estimate of a person's true long-term nutrient intake, but this estimate can be improved by using the average of multiple days of data for that person. The number of days needed has been discussed by Liu and co-workers (1978) and El Lozy (1983). Obviously, this number depends on both the degree of accuracy that is needed and the variability of the nutrient in question.

Beaton and colleagues (1979) have provided a simple formula that may be rearranged to calculate the number of days needed to estimate a person's true intake with a specified degree of error:

$$n = (Z_\alpha CV_w / D_o)^2 \qquad (3\text{-}3)$$

where: n = the number of days needed per person
Z_α = the normal deviate for the percentage of times the measured value should be within a specified limit
CV_w = the within-person coefficient of variation
D_o = the specified limit (as a percentage of long-term true intake)

The within-person coefficient of variation (CV_w) can be obtained from the analysis of variance on repeated days of dietary intake; the square root of the within-person variance is the within-person standard deviation and this value divided by the mean ($s/\bar{x} \times 100\%$) is the within-person coefficient of variation. Values for $\bar{x}$, and the within-person and between-person coefficients of variation for a variety of nutrients are shown in Table 3–5, based on data provided by the same 194 Boston-area women (Willett et al., 1985). Although these coefficients of variation may vary among different populations, they are remarkably similar to those reported for both men and women by Beaton and colleagues (1979), and thus probably are reasonable approximations for the general U.S. population.

Example: Suppose we wish to calculate the number of days needed to estimate a person's cholesterol intake to within 20 percent of their true mean 95 percent of the time. Thus $Z_\alpha = 1.96$ and CV_w from Table 3–5 is 62 percent:

$$n = (1.96 \times 62\%/20\%)^2 = 37 \text{ days}$$

Table 3–5. Means and within-person and between-person coefficients of variation for daily intake of selected nutrients.

		Coefficient of variation (%)			
		Unadjusted nutrients		Calorie-adjusted nutrients[a]	
Nutrient	Mean	*within-person*	*between-person*	*within-person*	*between-person*
Energy (kcal)	1620	27.0	19.3		
Protein (g)	68.3	32.9	16.4	25.0	14.0
Total fat (g)	68.6	38.4	22.6	27.3	14.1
Monunsaturated fat (g)	24.2	42.5	23.6	27.8	13.1
Polyunsaturated fat (g)	11.1	64.2	28.3	47.3	20.2
Cholesterol (mg)	311	62.2	23.8	61.5	24.1
Carbohydrate (g)	169.8	29.9	26.5	18.7	13.5
Sucrose (g)	46.9	60.3	45.3	50.1	29.1
Crude fiber (g)	3.27	44.3	31.5	31.0	23.0
Vitamin B_1 (mg)	1.08	40.8	23.3	36.6	17.5
Vitamin B_2 (mg)	1.43	39.1	23.1	35.3	18.4
Vitamin B_6 (mg)	0.85	51.9	31.1	21.4	12.8
Vitamin C (mg)	106.5	55.3	38.6	54.9	37.6
Vitamin A (IU)	5252	105.0	30.7	104.7	31.9
Iron (mg)	11.6	34.1	19.6	28.7	16.0
Calcium (mg)	616.9	41.9	28.3	36.1	21.7
Potassium (mg)	252.1	30.5	21.9	27.0	19.9

[a]Adjusted for caloric intake using regression analysis (see Chapter 12).
Data are based on four 1-week diet records completed by 194 U.S. women. (Willett et al., 1985).

Table 3–6. Number of repeated days needed per person for 95 percent of observed values to lie within specified percent of true mean

Nutrient	Within-person coefficient of variation	Number of days needed to lie within specified % of true means			
		10%	*20%*	*30%*	*40%*
Total fat	38.4	57	14	6	4
calorie-adjusted[a]	19.8	15	4	2	1
Cholesterol	62.2	149	37	17	9
calorie-adjusted[a]	61.5	145	36	16	9
Sucrose	60.3	140	35	16	9
calorie-adjusted[a]	50.1	96	24	11	6
Vitamin A	105.0	424	106	47	26
calorie-adjusted[a]	104.7	424	106	47	26

[a]Adjusted for total caloric intake using regression analysis.

Using such calculations, the number of days needed for the observed estimate of a person's intake to lie within a specified percentage of their true mean 95 percent of the time is given in Table 3–6. As can be appreciated, the days needed differ greatly for various nutrients. For many nutrients, obtaining a highly accurate estimate of individual intake by using repeated measurements is simply beyond practical possibilities in an epidemiologic study.

Implications for Developing Countries

The overwhelming importance of day-to-day variation in dietary intake has only recently been described formally in industrialized countries. Few analyses have been based on dietary intake within nonindustrialized populations. Conventional wisdom suggests that diets of poor populations in nonindustrialized areas are homogeneous, so that within-person variation may not be a serious consideration for epidemiologic studies. This issue deserves to be explored as 24-hour recalls and direct observations for a limited number of days are commonly employed in nonindustrialized countries.

Despite conventional wisdom it seems likely that important sources of variation in daily dietary intake may well exist in the developing world, even though the number of foods available may be limited. Where economic resources are severely restricted, food intake is strongly linked to income so that even small economic differences are directly reflected in diet. This linkage would tend to increase between-person variation. As in industrialized countries, individual preferences and cultural taboos also contribute to between-person differences in diet. Day-to-day variation may be particularly large in developing countries if expensive foods can be afforded only irregularly. For example, if meat is only eaten twice per week and caloric intake on other days is primarily obtained from a carbohydrate staple, then protein intake may appear high on some days and inadequate on others; that is, the within-person variance may be large. If preservation and transportation facilities are lacking, the influence of season may be much stronger than in industrialized countries, again increasing within-person variation.

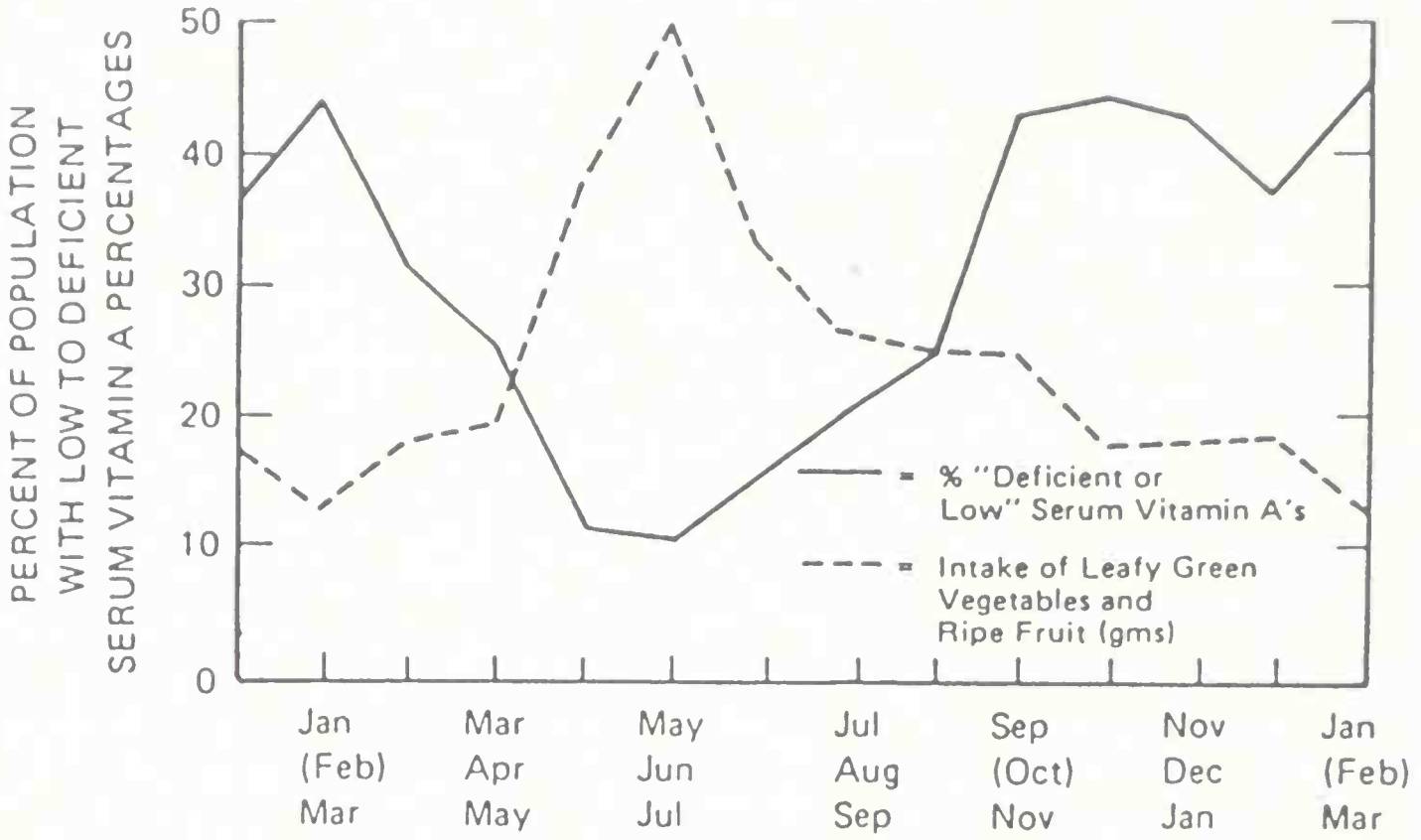


Figure 3–5. Seasonal variation of intake of leafy green vegetables and ripe fruit and of serum vitamin A in Bangladesh. (From Schaefer, 1981.)

Data from a national nutrition survey in Bangladesh that included over 25,000 persons (Schaefer, 1981) provided a striking example of seasonal variation (Fig. 3–5). From November through March the intake of leafy green vegetables and ripe fruit is very low and is temporally associated with a striking rise in the percentage of the population with low or deficient serum vitamin A levels (less than 20 μg/dl). A single measurement of diet in June would obviously provide a very misleading representation of usual carotene intake for the population as well as for most individuals.

The potentially large sources of between-person variation in diet in nonindustrialized countries may provide excellent opportunities for epidemiologic studies of the relation between diet and health. The similarly large potential for day-to-day variation means that careful, formal studies must be undertaken in such populations to measure the important components of variation, including seasonality, before embarking on any study based on a short-term measure of dietary intake.

Effects of Random Within-Person Variation on Measures of Associations in Epidemiologic Studies

Day-to-day variation in individual dietary intake has important implications for common measures of association in epidemiologic studies. If only one or a few days are measured, a subject's true long-term intake is likely to be misrepresented, that is, his or her dietary intake may be misclassified. This within-person variation, which may be regarded as random fluctuation above and below a person's true long-term average, can substantially distort correlation coefficients, regression coefficients, and relative risks. The general effect is to reduce the strength of associations.

The effect of within-person variation on the correlation between two normally distributed variables, x and y, has been discussed by Beaton and co-workers (1979), Liu and co-workers (1978), and Madansky (1959). Within-person variation (i.e., random error) in either x or y will cause the observed correlation between these variables to be attenuated, meaning lower than the true correlation between x and y. Specifically, let us consider the classical error model

$$x = \mu_x + \epsilon_x$$
$$y = \mu_y + \epsilon_y$$

where $\epsilon_x \sim N(0, \sigma^2)$, $\epsilon_y \sim N(0, \sigma^2)$. In this model, x and y represent observed values for one person based on one or more (n_x, n_y) repeated measurements. The underlying mean values are μ_x and μ_y, that is the mean values that would be observed if no measurement error were present (referred to as the true mean values), and ϵ_x and ϵ_y represent measurement error about the true mean values. If only a small number of measurements are made of x and y per individual, then the observed correlation (r_o, the correlation between x and y) is smaller than the true correlation (r_t, the correlation between μ_x and μ_y). The true correlation can also be interpreted as the correlation that would be observed if the average of many measurements were used to compute x and y for each person.

To estimate r_t (the true correlation), two methods have been proposed. The first is to collect a large number of replicate measurements for each individual and use the average of these to approximate the true value of x and y for that subject in computing the correlation coefficient. For example, Liu and colleagues (1979) have demonstrated that 14 24-hour collections of urine per subject are necessary to obtain an observed correlation between sodium excretion and another variable (say blood pressure), that is within 10 percent of the true correlation. Similarly, Beaton and co-workers (1979) and Sempos and co-workers (1985) have shown that, for most nutrients, many days of dietary intakes are necessary to avoid a major attenuation in the correlation between a nutrient and another factor. An alternative way to estimate the true correlation (r_t) is to make a small number of repeated measurements per subject and use knowledge of the within-person variance of x and/or y to estimate the true correlation. This approach is discussed in Chapter 12.

van Staveren and colleagues (1986a) have provided an example of the degree to which the use of a single 24-hour recall can reduce the true correlation between dietary intake and another variable. They first used the average of 19 24-hour recalls per person to compute the correlation between linoleic acid intake (expressed as the ratio of linoleic acid to saturated fatty acid intake) and the linoleic acid content of adipose tissue among 59 Dutch women. This correlation was 0.62. They then correlated each of the single days of linoleic acid intake alone, that is, one per person with the adipose linoleic acid levels; these correlations using individual 24-hour recalls were substantially lower, ranging from 0.14 to 0.50, with a median of 0.28.

Regression coefficients are attenuated by within-person variation in the independent variable. In an experimental feeding study, for example, it has been found that serum cholesterol increases 12 mg/100 ml for each increase of 100 mg/1000 kcal in dietary cholesterol (Mattson et. al., 1972). That is to say, the

slope of the regression line relating these variables is about 0.12 mg/100 ml change in serum cholesterol per milligram per 1000 kcal change in dietary cholesterol. By contrast, Beaton and colleagues (1979) used their data on the within-person variation of cholesterol intake to estimate that the observed regression coefficient would be only about 0.01, rather than 0.12, if a single day of cholesterol intake data were used to describe each subject's diet. The statistical relationship between the true and observed regression coefficients is discussed further in Chapter 12. In contrast with the effects of within-person error in the independent variable, random variation in the dependent variable does not systematically bias (i.e., attenuate) the regression coefficient. Within-person error in the dependent variable, however, increases the standard error of the regression coefficient.

The effect of random within-person variation on relative risk estimates is perhaps less obvious but no less important. Dietary factors are usually continuous variables; thus, in epidemiologic studies we are fundamentally comparing the distribution of a continuous variable among persons with a disease and without the disease. (In a prospective cohort study, the previous diets of subjects who became diseased would be compared with the previous diets of the total group or of those who remained free of disease.)

These distributions can be portrayed as in Figure 3–6A where it is assumed that diet is measured without random error, that is, there is no within-person variation. To compute a relative risk, we need to specify a cut-point to define high and low intakes. This point is sometimes rather arbitrary and often several cut-points are used. For simplicity in this example we have used a cut-point (x_n) corresponding to 1 standard deviation (SD) above the mean of the noncases. Let us assume that the distributions for both groups are normal and that the mean nutrient intake for cases is 0.5 SD above the control value.

The proportions of subjects in each of the cells of a 2×2 table are then described as the areas defined by these distributions, these values being obtained from a table of normal deviates:

Nutrient intake		
	High	Low
Cases	$a = 0.31$	$b = 0.69$
Noncases	$c = 0.16$	$d = 0.84$

where

a = area proportional to the number of cases above cut-point (x_n)
b = area proportional to the number of cases below cut-point (x_n)
c = area proportional to the number of non-cases above cut-point (x_n)
d = area proportional to the number of non-cases below cut-point (x_n)

The odds ratio is then computed as:

$$\frac{a/b}{c/d} = \frac{0.31/0.69}{0.16/0.84} = 2.36$$

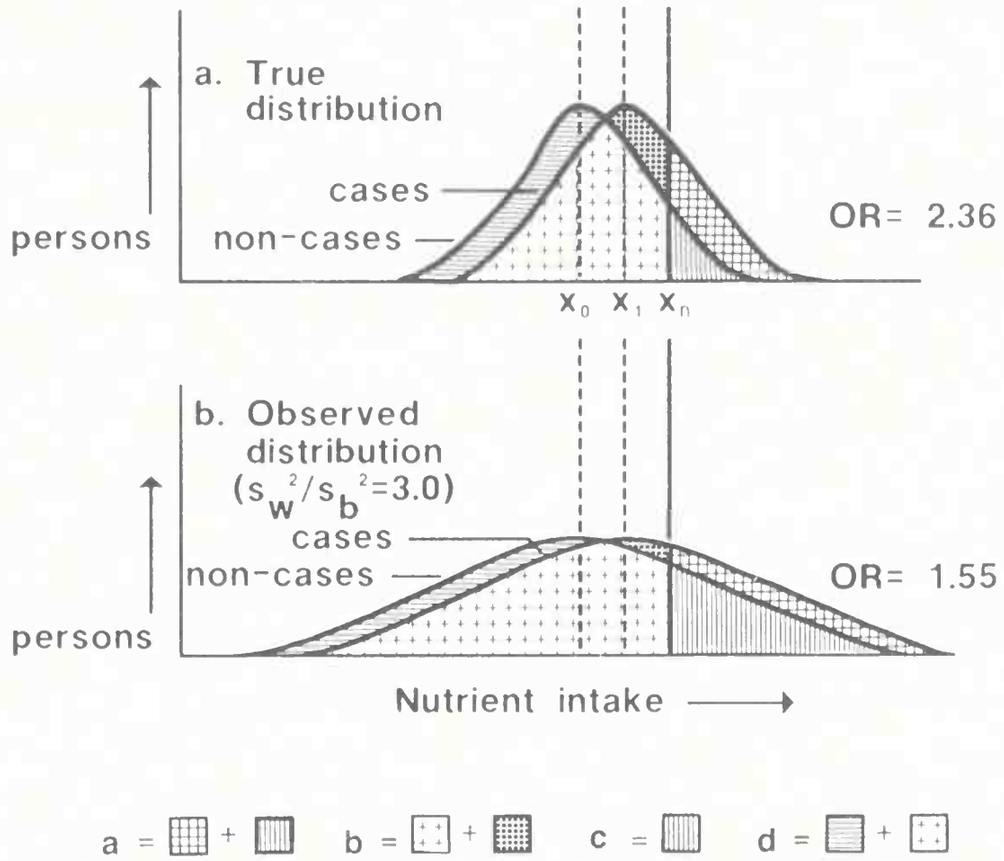


Figure 3–6. **A.** Hypothetical true distributions for noncases (mean nutrient = x_0) and cases (mean nutrient intake = x_1, which is 0.5 standard deviation larger than x_0). x_n is an arbitrary cut-point 1.0 standard deviation above x_0: values higher than x_n are considered high, values below x_n, are considered low. **B.** Observed distributions for the same cases and noncases based on a single observation per subject when the within-to-between variance ratio for the nutrient = 3.0. The observed SD is now twice the true SD.

Now let us assume that the single measurement of intake is subject to random error. As described previously, the distributions for both cases and noncases are wider: the standard deviations are increased. In this example, assume that the ratio of within-person to between-person variance is 3.0, which is realistic for many nutrients (see Table 3–4). Because the between-person variance is unchanged, as the subjects are the same, this implies that the observed standard deviation is twice as large as the true standard deviation, s_b:

$$\text{observed SD} = (s_b^2 + 3s_b^2)^{1/2} = 2s_b$$

These broader observed distributions are shown in Figure 3–6B; it is obvious that the case and noncase distributions are now less distinct. The observed areas corresponding to the cells of a 2 × 2 table can be recalculated from a table of normal deviates (note that the same change in nutrient intake corresponding to 1 normal deviate for the true distribution is now 0.5 of a normal deviate for the observed distribution):

Nutrient intake		
	High	Low
Cases	$a' = 0.41$	$b' = 0.59$
Noncases	$c' = 0.31$	$d' = 0.69$

OR = 1.55

It is apparent that the observed odds ratio has been attenuated as it is now considerably closer to the null value of 1.0.

SUMMARY

The day-to-day variation in nutrient intake among free-living subjects has consistently proved to be large, although the magnitude varies according to nutrient. Measurements of dietary intake based on a single or small number of 24-hour recalls per subject may provide a reasonable (unbiased) estimate of the mean for a group, but the standard deviation will be greatly overestimated. Furthermore, measurements of association in epidemiologic studies, such as correlation and regression coefficients and relative risks, are substantially weakened, possibly to the point of being undetectable.

REFERENCES

Beaton, G. H., J. Milner, and P. Corey, et al. (1979). Sources of variance in 24-hour dietary recall data: Implications for nutrition study, design and interpretation. *Am. J. Clin. Nutr. 32,* 2546–2559.

Beaton, G. H., J. Milner, and V. McGuire, et al. (1983). Sources of variance in 24-hour dietary recall data: Implications for nutrition study, design and interpretation. Carbohydrate sources, vitamins, and minerals. *Am. J. Clin. Nutr. 37,* 986–995.

Brown, K. H., R. E. Black, and S. Becker (1982). Seasonal changes in nutritional status and the prevalence of malnutrition in a longitudinal study of young children in rural Bangladesh. *Am. J. Clin. Nutr. 36,* 303–313.

Dalvit, S. P. (1981). The effect of the menstrual cycle on patterns of food intake. *Am. J. Clin. Nutr. 34,* 1811–1815.

El Lozy, M. (1983). Dietary variability and its impact on nutritional epidemiology. *J. Chron. Dis. 36,* 237–249.

Hunt, W. C., A. G. Leonard, P. J. Garry, and J. S. Goodwin (1983). Components of variance in dietary data for an elderly population. *Nutrition Research 3,* 433–444.

Liu, K., J. Stamler, and A. Dyer, et al. (1978). Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J. Chron. Dis. 31,* 399–418.

Liu, K., R. Cooper, and J. McKeever, et al. (1979). Assessments of the association between habitual salt intake and high blood pressure: Methodological problems. *Am. J. Epidemiol. 110,* 219–226.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Am. Statis. Assoc. J. 54,* 173–205.

Mattson, F. H., B. A. Erickson, and A. M. Kligman (1972). Effect of dietary cholesterol on serum cholesterol in man. *Am. J. Clin. Nutr. 25,* 589–594.

Morgan, K. J., S. R. Johnson, and B. Goungetas (1987). Variability of food intakes. An analysis of a 12-day data series using persistence measures. *Am. J. Epidemiol. 126,* 326–35.

Rush, D. and A. R. Kristal (1982). Methodologic studies during pregnancy: the reliability of the 24-hour dietary recall. *Am. J. Clin. Nutr. 35,* 1259–1268.

Schaefer, A. E. (1981). Can nutritional status be determined from consumption or other measures? In *Assessing Changing Food Consumption Patterns.* Washington, D.C.: National Academy Press.

Sempos, C. T., N. E. Johnson, E. L. Smith, and C. Gilligon (1985). Effects of intra-individual and inter-individual variation in repeated dietary records. *Am. J. Epidemiol. 121,* 120–130.

U.S. Dept. of Health, Education, and Welfare. Public Health Service, National Center for Health Statistics. *Dietary Intake Source Data, 1971–74.* DHEW Pub No. (PHS)79–1221.

van Staveren, W. A., P. Duerenberg, M. B. Katan, J. Burema, L. C. deGroot, and M. D. Hoffmans (1986a). Validity of the fatty acid composition of subcutaneous fat tissue microbiopsies as an estimate of the long-term average fatty acid composition of the diet of separate individuals. *Am. J. Epidemiol. 123,* 455–463.

van Staveren, W. A., P. Duerenberg, J. Burema, L. C. deGroot, and J. G. Hautvast (1986b). Seasonal variation in food intake, pattern of physical activity and change in body weight in a group of young adult Dutch women consuming self-selected diets. *Int. J. Obes. 10,* 133–145.

Willett, W. C., L. Sampson, M. J. Stampfer, et al. (1985). Reproducibility and validity of a semi-quantitative food frequency questionnaire. *Am. J. Epidemiol. 122,* 51–65.

Zeigler, R. G., H. B. 3d Wilcox, T. J. Mason, J. S. Bill, and P. W. Virgo (1987). Seasonal variation in intake of carotenoids and vegetables and fruits among white men in New Jersey. *Am. J. Clin. Nutr. 45,* 107–114.

4

Short-Term Dietary Recall and Recording Methods

JELIA C. WITSCHI

In 1938, Burke and Stuart had mothers keep a 24-hour, meal-by-meal food intake record for their children who were enrolled in the Center for Research in Child Health and Development at Harvard School of Public Health. The mothers were then interviewed (as a cross-check) regarding their child's previous 24-hour food intake. This study was among the earliest descriptions of dietary intake methodology and incorporated two of the approaches in common use today: the 24-hour recall and the dietary record. In this chapter, short-term recall and dietary recording methods are described and information regarding their validity is reviewed. In particular, the limitations and utility of these approaches in the context of epidemiologic studies is examined.

METHODOLOGY

Twenty-Four-Hour Recall

The 24-hour recall, pioneered by Burke (1947), McHenry (1939), and Kruse and colleagues (1940), is the most widely used dietary assessment method. The Nationwide Food Consumption Surveys (NFCS) (Pao et al., 1985), Ten-State Nutrition Survey (1972), Health and Nutrition Examination Surveys (HANES) (Carroll et al., 1983; Abraham et al., 1979), and other large studies, such as the Multiple Risk Factor Intervention Trial (Tillotson et al., 1981), have used it as the primary dietary assessment method, either alone or in combination with other methods. Interviews are conducted by nutritionists or trained interviewers using standardized forms and visual references. The method is relatively rapid, requiring 10 to 20 minutes for trained interviewers, and is, therefore, attractive from a budgetary perspective.

As its name implies, the 24-hour recall is an attempt to define and quantify food intake during a specific day just before the interview. Interviews may begin at the present and look backward, or begin at breakfast of the proceeding whole day and work forward. The success of 24-hour recalls depends on the memory,

cooperation, and communication ability of the subject, and on the skill of the interviewer.

The use of key questions and memory aids is crucial to obtaining accurate information about types and quantities of foods eaten. Discussions may center around the main dish in a meal, relate food intake to particular events or activities, or focus on nonmeal eating such as snacking while watching television. Probing questions may be used to elicit more information, and checklists can be helpful reminders of easily forgotten foods. Food models are often used to represent portion sizes. In a study of 30 couples, Moore and co-workers (1967) found that estimations of food intake were more accurate when food models were used. Similarly, Guthrie (1984) determined that the ability of young adults to select and immediately describe predetermined quantities of food without the aid of measuring devices was poor.

Twenty-four-hour recalls conducted by telephone have been described by several investigators (Krantzler et al., 1982a; Posner et al., 1982; Schucker, 1982). Posner and colleagues (1982) mailed a two-dimensional visual aid to participants to assist in portion size estimation. Morgan and colleagues (1987) compared use of telephone with personal interview dietary assessments, and concluded that the telephone interview provided comparable data with less effort and cost. Response rates were 71 to 81 percent throughout four contacts (compared with 72 to 83 percent for personal interviews) and the total time required for the telephone technique averaged less than 2 hours, which was shorter than for personal interviews.

Dietary Records

Dietary records or food diaries are detailed descriptions of types and amounts of foods and beverages consumed, meal by meal, over a prescribed period, usually 3 to 7 days. The record or diary may be a special form or booklet, lined or unlined, with pages that are either blank or that contain suggested categories of foods for each day. In some applications, foods are weighed or measured using specified procedures. Regardless of the specific details, dietary recording places a substantial burden on the subject, which limits the use of the method to persons who are literate and highly motivated (Mahalko et al., 1985).

Recording procedures, particularly relating to the detailed descriptions of types and quantities of foods, must be taught to participants. In general, immediate recording is desirable to minimize memory loss. The act of recording may raise awareness of food to the point of altering eating behavior (Mahalko et al., 1985). This effect is undesirable when the records are being used to represent a person's usual intake. In dietary intervention programs, however, this heightened awareness can be used to advantage; thus diet records are frequently used as a teaching device.

SOURCES OF ERROR IN SHORT-TERM RECALLS AND RECORDS

Three major sources of error potentially affect all dietary survey methodologies: the respondent (recalls) or recorder (records); the interviewer (recalls) or inter-

preter or reviewer (records); and the database from which nutrient intakes are computed.

Respondent and Recorder Errors

The ability of respondents to provide reliable dietary data depends on their motivation, awareness of food intake, memory, and communication skills. In a longitudinal study of adolescents, Hackett and co-workers (1985) found that survey fatigue, learning effect, season of the year, day of the week, and the subjects' knowledge of the study's purpose influenced the measurement of food intake. Kim and colleagues (1984a) indicated that longer study periods and continual contact with the investigators contributed to the accuracy of reports. In some instances subjects may be unwilling to report accurately; Paul and co-workers (1963) found that "the men often conceded that they were embarrassed to list in some cases all that they really had consumed."

The perception of foods eaten, both in type and amount, is obviously central to the success of dietary recalls. Compared with weighing, the use of household measures to determine food quantities in interviews or dietary records, which are then converted to gram weights, clearly results in a loss of precision (Marr, 1971). Casual estimates of food quantity (large or small) are usually very inaccurate and may produce the largest errors in the estimate of portion size. Guthrie (1984) observed major discrepancies among 147 young adults between perceived standard portions and amounts of food regarded as typical portions by nutritionists. When asked to describe food portions in terms of common measures, these subjects both overestimated and underestimated predetermined servings of breakfast and luncheon foods. Amounts of butter, salad dressing, cereal, and salads were overestimated by more than 51 percent up to two-thirds of the time; intake of salad dressing, butter, sugar, and salad were underestimated by at least 51 percent up to one-fourth of the time.

Lansky and Brownell (1982) found that differences in estimated quantities account for substantial errors in self-reported food intake. In a study of 30 obese female applicants to a weight reduction program, each subject was shown 10 commonly used foods and was asked to estimate quantities and caloric equivalents of each food. Quantities were overestimated by an average of 63.9 percent; the errors ranged from 6 percent (cola) to 260 percent (potato chips).

Memory curves demonstrate that the rate of forgetting is a negatively accelerating phenomenon (Becker et al., 1960), and the memory of food intake fades rapidly as the number of items increase. Selective suppression and distortion in memory of specific foods may affect dietary assessment. Unpleasant events (such as perceived poor meals) may be relegated to the subconscious and occurrences that do not correspond to expectations may be suppressed. Socially acceptable events and foods may be better remembered and reported than those that are less acceptable. Worsley and colleagues (1984) found that reported intakes of certain foods (fresh fruits, vegetables, and sweet foods) are susceptible to social desirability biases. Youland and Engle (1976) reported that the largest source of error in HANES was the uncertainty of subjects about foods consumed on the recall day.

The effects of age, sex, and environment on memory of foods consumed were examined by Campbell and Dodds (1967) in a population of 200 elderly institutionalized persons over 65 years age, and 100 people 20 to 40 years old. After a 24-hour recall, memory loss was evaluated by extensive probing including reading back the recall and, for institutionalized persons, reading the previous day's menu. An additional 35 percent of calories were recalled from institutionalized men after probing; 28 percent from institutionalized women; and 21 percent from younger men, 12 percent from younger women. The researchers concluded that recall was better by women, by younger respondents, and by older respondents living at home. Pekkarinen and co-workers (1967) found that, among rural Finnish families, nutrient intakes could be overestimated or underestimated by a factor of up to three when a 24-hour recall was compared with a more-or-less concurrent precise weighing technique, a situation that might be expected to sensitize participants to report more accurately.

It has been suggested that obese subjects tend to underreport food intake; however, this has been difficult to document without an objective measure of true dietary intake suitable for use among free-living persons. The recently available doubly labeled water method, which uses stable isotopes of hydrogen and oxygen, can provide such an assessment. Using this method, Prentice and colleagues (1986) compared total daily energy expenditures of 9 lean and 13 obese postpartum women with their dietary records. Isotopically measured total energy expenditures of obese subjects were 28 percent higher than lean controls (2445 vs. 1911 kcal). Self-recorded mean energy intakes (obtained by weighing all food and drink for 7 to 14 days), however, were underestimated by 837 kcal/day among obese women, but were accurately estimated (within 34 kcal) by lean women.

The process of collecting dietary data may itself affect eating behavior (Marr, 1971). Kim and colleagues (1984b) determined that 28 of 29 subjects participating in a year-long study during which continuous daily diet records were submitted, decreased their energy intakes from 1.1 to 32.3 percent (mean decreases of 12.9 percent from yearly mean) during periods of duplicate food collections. Greater decreases in specific nutrients, particularly protein and fat, suggested that subjects selectively reduced their intake of more expensive foods during the periods of duplicate collections, perhaps due to the need to purchase additional food within constrained budgets.

Errors may be introduced by failure to report the extent of eating away from home as the memory of these foods may be less accurate than for those eaten at home. Quantities consumed are more difficult to estimate and specific details of ingredients and food preparation may be unknown. The frequency of eating away from home continues to rise and its importance in relation to the accuracy of dietary recall has not been extensively studied.

Noncompliers to study protocols are likely to be different from persons who comply (Morgan et al., 1971), so a low dropout rate may be important to maintain generalizability. Gersovitz and colleagues (1978) noted that as food recordings progressed toward the seventh day, bias was introduced because of dropouts and decreased quality of records. The extent of bias contributed by nonresponders to a population study is difficult to document.

Interviewer and Reviewer Errors

Well-trained interviewers may obtain comparable results. Frank and co-workers (1977) found that the reliability of the 24-hour recall was improved by training interviewers using a detailed protocol. Standardized procedures for probing and coding were thought to be particularly important (Frank et al., 1984). In a study of sources of variance in 24-hour recalls, standardization and certification of nutritionist interviewers was credited by Beaton and colleagues (1979) with having eliminated interviewer differences as an important source of variation. Ernst and colleagues (1980) have described the components of a standardized system for collecting and processing dietary information for use in large studies. Efforts to develop objective, standardized coding procedures for portion sizes, gram-weight conversions, food preparation factors, edible portions, and nutrient absorptions have been reported by Hankin and Hunemann (1967) and Houser and colleagues (1969).

Errors may be introduced by the interviewer because of behavioral factors, such as the manner of asking questions. Information may vary depending on whether probing is complete and detailed, or vague and general. Failure to probe for details was a frequent source of error among interviews in HANES (Youland and Engle, 1976). Other sources of error that are very difficult to document may be introduced by the interviewer. Mannerisms, various gestures, or nonverbal body language may be distracting or suggestive of preferred or "appropriate" responses. The interviewer may influence reporting by conveying personal opinions through reactions to a subject's responses. Social status and attitudes as well as personal food habits of the interviewer may seriously influence reporting. A feeling of rapport and an unbiased attitude toward the respondent are essential.

Clerical or coding errors may seriously affect dietary data. Youland and Engle (1976) developed a test for interviewers in HANES to determine the degree to which they were following protocol and found that the most frequent errors were use of wrong food codes and transformed numbers in coding. Adelman and colleagues (1983) compared nutrient data calculations by two experienced users of a single computerized nutrient database using 3-day food records supplied by 10 volunteers. No statistically significant differences in means were found; however, the standard deviation of the difference between observers was approximately 165 kcal/day for energy to 1.5 mg/day for iron. These sources of error can potentially be reduced by interactive, computerized data-entry systems that eliminate numerical codes for foods (see Chapter 2).

Nutrient Database Errors

Nutrient databases cannot provide an exact measure of an individual's nutrient intake (see Chapter 2). The nutrient data are derived from the analysis of food samples from various sources and, as such, have limited accuracy when applied in a particular case (Acheson et al., 1980; Jacobs et al., 1985). Although databases have major limitations, they are, of course, far easier than direct laboratory analysis. Jacobs and colleagues (1985) submitted 24-hour records from 54

middle-aged men to three different databases for nutrient computation and found that differences were due principally to systematic coding and standardization procedures. Systems requiring hand calculations in coding procedures (e.g., converting servings to gram weights) exhibited the greatest errors. For example, there were significant differences in the computation of mean alcohol intake (22.3 to 27.2 g, $p < 0.01$), percent of calories from alcohol (7.0 to 8.9%, $p < 0.01$), and percent of calories from carbohydrate (34.7 to 37.4%, $p < 0.01$). These differences were believed to be related primarily to the way in which calories from alcohol were computed in one system. A difference was also seen in distribution of fat among the systems; one system computed 1.4 percent more calories from polyunsaturated fat.

In recent years, most nutrient database systems have improved, but substantial differences still exist among them. Hoover (1983) compared analyses of dietary records among 8, 7, and 11 different computerized nutrient databases in 3 successive years. Differences between highest and lowest values were greatest in the first year (25% for carbohydrate to 66% for protein); differences were reduced by the third year (7% for protein to 45% for fat); the improvements were believed due to improved quantification of portion sizes. Differences in the computation of fat intake (60% in year 1, 66% in year 2, and 45% year 3), were attributed to variation in the nutrient databases.

Unavailable data for new foods or formulations and culture-specific foods represent error sources of unknown dimensions. Little research has been done regarding the optimal strategy for dealing with incomplete food composition values, but it is reasonable to assume that there are significant effects (see Chapter 2). Major disparities exist in knowledge of nutrient composition of foods; for example, information on amino acid content is complete for most foods, whereas values for simple sugars are generally unavailable. The USDA National Nutrient Data Bank has been described (Hepburn, 1982) as a significant development that will lead to improved knowledge of nutrient composition of foods through nationwide sampling, accurate and efficient analysis, improvement in data handling and processing, and the standardized application of computerized database information to surveys.

REPRODUCIBILITY AND VALIDITY

Reproducibility

The concept of reproducibility is complex when considering short-term dietary records or recalls as diet itself varies dramatically from day to day. Because the issue of temporal variation is discussed in detail in Chapter 3, reproducibility is considered here in reference to simultaneous measurements of the same individuals in the same situation and time frame (Becker et al., 1960). Such information is obviously difficult to obtain as identical interview situations are never achieved in practice. Thus it is difficult to distinguish real changes in diet from errors in measurement (Rasanen, 1979). Because each assessment affects subsequent assessments, repeat interviews should ideally not be so close as to influence answers or so distant as to affect memory. Because these conditions are

difficult to achieve, few data are available on the reproducibility of these short-term methods. In one of the few such studies, Frank and co-workers (1984) examined recording practices of two independent interviewers in the Bogalusa Heart Study among a group of 18 10-year-olds. For the same 24-hour period the two interviewers provided 71 to 82 percent agreement on food names, 66 to 100 percent agreement on food identification numbers, and 85 to 91 percent agreement on food quantities. The greatest differences in quantitation appeared for meats, liquids, and sweets.

Validity

To establish validity, by definition, requires a true, accurate measure of dietary intake (see Chapter 6). Liu and colleagues (1978) have noted the difficulty of determining the validity of dietary data of free-living people due to limitations in measurement techniques. Although it is possible to structure research procedures that can measure the validity of the 24-hour recall, validation studies of dietary records are more difficult, and few have been reported.

Attempts to measure validity inevitably involve a comparison of survey instruments or procedures. Although there may be preferred methods for particular purposes, there is, unfortunately, no ideal method of dietary assessment against which all other methods may be calibrated (Beaton et al., 1983). Validation of a method must be viewed in relation to the purposes for which the research is being conducted; for example, whether it requires comparison of groups, individuals, or only the identification of individuals at the extremes of a distribution. Five general approaches to validate dietary methods are: (1) observation of intake, (2) weighing food before selection and consumption, (3) comparing two approaches of reporting intake, (4) laboratory analysis of duplicate meals or food portions, and (5) biochemical determinations of a physiologic variable related to a specific nutrient. Unobstrusive observation of actual food intake by subjects has been used for the validation of the dietary recall method. Karvetti and Knuts (1985) observed the actual intake of 140 subjects and later interviewed them by 24-hour recall (Fig. 4–1). They found that subjects erroneously recalled foods that were not actually eaten, ranging from 5 to 29 percent of specific foods and omitted foods that were actually eaten, ranging from 4 (fish) to 50 percent (cooked vegetables). The correlations between nutrients calculated from observed intake with those based on recalled intake ranged from 0.58 to 0.74. Madden and colleagues (1976) compared observed intakes of noon meals consumed by 76 elderly persons with recall of that meal within the following 24 hours. Significant differences in means were observed only for calories, with recall being about 10 percent lower. Correlations between nutrient intakes assessed by the two methods ranged from 0.28 for protein to 0.87 for ascorbic acid. Carter and co-workers (1981) studied 28 children, aged 10 through 12, and found large and significant differences between recalled mean and observed mean intake for kilocalories and protein. For kilocalories, the observed mean was 2348, whereas the recalled mean was 1896 ($p < 0.002$); for protein, the observed mean was 82 g and the recalled mean 66 g ($p < 0.004$).

The technique of weighing foods before consumption and administration of

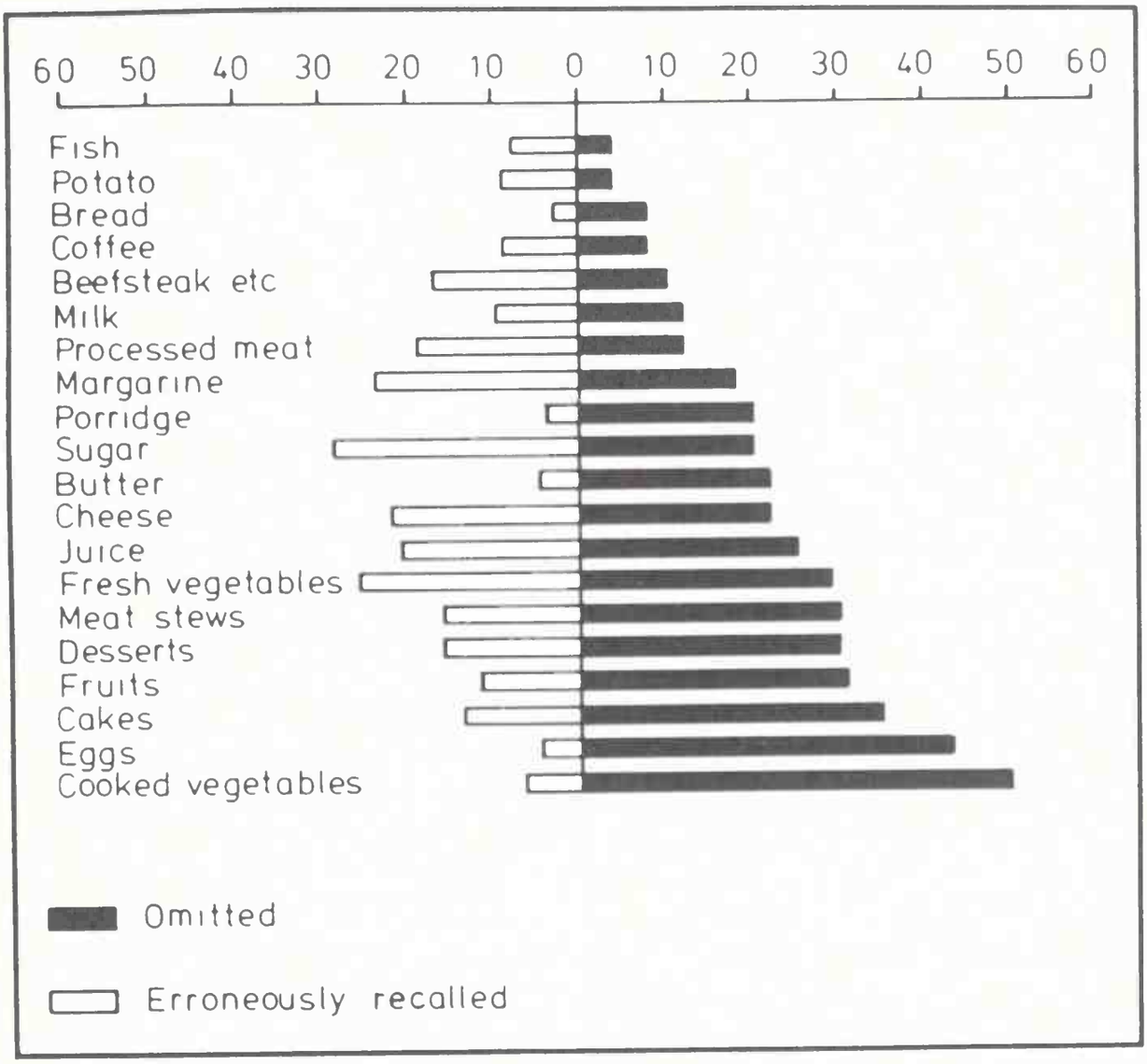


Figure 4–1. Percentage of times that 140 persons either failed to report or erroneously reported specific foods when 24-hour recalls were compared to observed intake. (From Karvetti and Knuts, 1985; reproduced with permission.)

recall can evaluate accuracy in quantitation as well as memory of foods eaten. Greger and Etnyre (1978) studied 32 adolescent women for 30 days in a metabolic ward. Twenty-four-hour recalls were conducted on the second and fifth days of the study with 17 and 15 participants, respectively. On the first recall day (Table 4–1), these girls recalled substantially less than actual intakes of vitamins A, thiamin, riboflavin, niacin, ascorbic acid, and iron; on the second recall day, they significantly overestimated niacin, riboflavin, and iron. Foods omitted from the recalls (Table 4–2) ranged from 11.8 percent of beverage servings to 36.6 percent of other solid foods, whereas foods added or misidentified ranged up to 9.4 percent of milk and milk products. Schnakenberg and colleagues (1981) weighed and observed food intake of 62 military personnel to validate a diary/interview technique and found that food item selection agreed 80 percent of the time and nutrients were underestimated by about 10 percent using the diary/interview methods, with the exception of protein, fat, and niacin, which did not differ. Fourteen percent of food items were observed but not reported to have been consumed and 6 percent were reported but not observed.

A common validation procedure is to compare the mean values for a group using two different dietary assessment methods. Mean nutrient intakes based on

Table 4–1. Mean nutrient intakes from food: Actual (weighed) versus recalled intakes

	Day 1 (n = 17)		Day 2 (n = 15)	
Nutrient	*Actual*	*Recall*	*Actual*	*Recall*
Energy (kcal)	2172	2000	2474	2429
Protein (g)	44	51	46	54
Vitamin A (IU)	7663	3377[a]	7602	8692
Thiamin (mg)	2.3	1.4[a]	1.8	2.1
Riboflavin (mg)	2.7	1.6[a]	1.8	2.5[a]
Niacin (mg)	29	16[a]	20	26[b]
Ascorbic acid (mg)	116	78[a]	201	174
Calcium (mg)	391	418	585	713
Iron (mg)	25.5	14[a]	18.6	23[b]
Zinc (mg)	6.7	8	12.2	11

[a]Significant differences ($p < 0.001$) between actual and recall.
[b]Significant differences ($p < 0.05$) between actual and recall.
From Greger and Etnyre, 1978.

24-hour recalls were compared with those from 1-day food records collected 2 days later by 2667 elderly persons (Fanelli and Stevenhagen, 1986). Remarkably similar estimates of mean energy and nutrient intakes were found for both men and women (Table 4–3). Rasanen (1979) compared 24-hour recalls and dietary histories taken at the same interview from 741 children, ages 5 to 13. Mean nutrient values obtained by the history method were significantly higher by 32 percent (carbohydrate and calcium) to 56 percent (vitamin A) than those from 24-hour recalls. Correlations ranged from 0.20 for vitamin A to 0.50 for energy.

The validation of diet records that measure intake for longer periods of time may be extremely difficult or impractical. Even short-term dietary records are difficult to validate. In practice, dietitians normally check dietary records with clients to complete inadequate descriptions and quantitation or to inquire about possible forgotten or ignored food items. Unless recording is done immediately after each meal, very limited probing is possible because the memory span for food items and quantities is so limited. The use of multiple options for units of portion sizes, such as interchanging volumes and weights, within the same record can increase imprecision.

Table 4–2. Frequency with which foods were omitted, added, or misidentified in dietary recalls

	Actual diet (servings/person during 2 days)	Percent of servings	
Food group		*Omitted*	*Added/misidentified*
Milk/milk products	2	12.5	9.4
Meats or alternate	5	18.7	5.1
Fruits and vegetables	14	29.4	6.8
Breads and cereals	9	26.3	2.5
Other beverages	6	11.8	1.0
Other solid foods	12	36.6	1.4

From Greger and Etnyre, 1978.

Table 4–3. Comparison of mean intakes of energy and 14 nutrients from 24-hour recall and 1-day record for older men and women

	Men				Women			
	65–74 yr (n = 686)		*≥75 yr (n = 326)*		*65–74 yr (n = 1057)*		*≥75 yr (n = 598)*	
Nutrient	Recall	Record	Recall	Record	Recall	Record	Recall	Record
Energy (kcal)	1,929	1,924	1,840	1,878	1,430	1,439	1,381	1,453
Protein (g)	78.8	79.9	74.0	76.3	60.4	61.7	55.6	59.3[a]
Fat (g)	89.1	87.9	84.9	85.0	63.2	63.4	59.5	63.3
Carbohydrate (g)	203.3	201.6	196.7	201.8	157.2	157.5	158.3	164.6
Calcium (mg)	708	715	696	717	556	560	576	612
Iron (mg)	14.2	14.3	14.0	14.0	10.6	10.9	10.4	10.9
Magnesium (mg)	284	284	266	272	226	226	221	230
Phosphorus (mg)	1,209	1,216	1,144	1,174	913	927	890	945
Vitamin A (IU)	6,437	7,967	6,311	7,455	6,346	6,486	5,848	6,581
Thiamin (mg)	1.38	1.39	1.36	1.33	1.06	1.05	1.04	1.08
Riboflavin (mg)	1.78	1.88	1.72	1.80	1.38	1.42	1.37	1.48
Niacin (mg)	20.19	20.51	18.61	19.53	15.75	15.96	14.45	15.33
Vitamin B_6 (mg)	1.55	1.60	1.50	1.54	1.24	1.30	1.16	1.25
Vitamin B_{12} (μg)	5.04	7.11	4.52	6.04	4.01	4.81	3.80	4.40
Vitamin C (mg)	88	95	86	94	94	93	81	88

[a]Significant difference ($p < 0.001$) between recall and record.
From Fanelli and Stevenhagen, 1986.

Laboratory analysis of duplicate meals or food samples has been used as a validation technique. Witschi and colleagues (1985) used duplicate 24-hour food collections to validate the use of simultaneous 1-day food diaries in a subsample of 24 high school students. A correlation coefficient of 0.75 was observed between sodium values computed from food diaries and laboratory analysis of sodium in duplicate food collections.

Because of reported caloric decreases and apparently contradictory increases in body weight over the same time in national surveys (NFCS and HANES I and II), Mertz and Kelsay (1984) designed and conducted a comprehensive 1-year dietary intake study of 13 men and 16 women, ages 19 to 50 years. Participants submitted daily dietary records and four 1-week duplicate food collections spaced throughout the year. Miles and co-workers (1984) analyzed these duplicate food collections for energy, which was compared with kilocalories calculated from food tables. Calculated mean daily kilocalories for men were 2386, whereas analyzed mean kilocalories were 2327, a difference of 59 (2.5%); for women, 1638 kilocalories were calculated from food tables and 1685 were measured by analysis, a difference of −47 (−2.9%).

Investigators may indirectly validate dietary intake methods by comparing assessments with biochemical measures that would normally be expected to covary with dietary variables (Mahalko et al., 1985). For example, Caggiula and colleagues (1985) obtained 1-day food records and, on the same day, 24-hour urine collections from 55 adults to evaluate whether food records were appropriate for calculating sodium and potassium intake. They determined that for sodium, mean calculated dietary intake was 113 mEq and laboratory determi-

nation of average urinary excretion was 149.4 mEq; the correlation between the measurements was 0.53. For potassium, mean calculated dietary intake was 67 mEq and the mean laboratory determination of urinary excretion was 65.8 mEq; the correlation between measurements was 0.58. The authors concluded that the food records gave good group estimates of the excretion of sodium and potassium. Witschi and colleagues (1985) compared urinary sodium excretions with sodium determinations from 24-hour food collections and obtained a correlation of 0.58 for 24-hour urinary sodium (made on the same day), and a correlation of 0.33 for overnight urinary sodium (after food collection).

Bingham and Cummings (1985) measured 24-hour urinary and fecal nitrogen excretion and assessed protein intake by dietary records and analysis of replicate meals for three women and five men 24 to 57 years of age who were participating in a metabolic ward experiment. During the 28-day study, they each ate foods typical of their previously assessed usual diet. Nitrogen intake averaged 15.95 g/day by calculation from the diet record, 15.90 g/day by analysis of replicate meals, and 14.92 g/day by urinary and fecal excretion. The correlation between individual 28-day averages of calculated versus analyzed nitrogen intakes was 0.97.

Mahalko and colleagues (1985) obtained 7-day food records for 54 adults and observed a positive correlation ($r = 0.34$) between dietary iron and serum ferritin among unsupplemented subjects. Plasma ascorbate correlated with dietary intake examined by 7-day food records ($r = 0.34$) and with ascorbate supplements ($r = 0.53$). van Staveren and co-workers (1986) performed fat biopsies and also estimated diet in 59 Dutch women by taking the mean of 19 24-hour recalls over a period of 18 months and found highly significant correlations between linoleic acid content of fat tissue and in the diet ($r = 0.70$).

The interpretation of data from validation studies such as these is considered further in Chapter 6. In general, the validity of 24-hour recalls and diet records is likely to vary depending on the study population, the skill of the investigator, and the nutrients being measured.

LIMITATIONS OF SHORT-TERM RECALLS AND RECORDS

The choice of a dietary evaluation method depends on the objectives of the study; the degree of accuracy and the type of data needed; the skill of the professional and support staff; and the size, ability, and cooperation of the study population. The limitations of short-term recalls and records, as well as their appropriate place in nutritional epidemiology, is considered here.

Limitations of 24-hour Recalls

The most fundamental limitation of the 24-hour recall method is that dietary intake from day to day is highly variable (see Chapter 3). Any given 24-hour period is usually not typical or representative of long-term intake and may seriously misrepresent actual nutrient intake for an individual. In addition, certain days may be more representative of usual intake than others. Beaton and co-

workers (1979) studied 24-hour recalls and found a significant and consistent day-of-the-week effect for absolute nutrient intakes among women, with intakes on weekends being higher. Balogh and colleagues (1971) indicated that because of considerable variation in intake between days, the recalled day should be a "representative" day; unfortunately, such a day is difficult or impossible to identify.

Even for the day under consideration, evidence suggests that 24-hour recalls are usually not completely accurate. Foods actually eaten are often not remembered and reported foods are frequently not those consumed; selective recall appears to be influenced by perceptions of desirability. Quantities of foods eaten are easily forgotten; large intakes tend to be underreported and small intakes tend to be overreported.

Twenty-four-hour recalls are not applicable to most epidemiologic studies that require an assessment of long-term food consumption. Moreover, in case-control studies, past diet is of interest rather than current diet, which is provided by short-term recalls. In addition, the collection and processing of recall information is labor-intensive and subject to error. Bias may be introduced through the necessity to probe for possible incomplete information, the interpretation required to use the recalled information, and by the need to discard data that are incomplete, questionable, or considered unusable.

Limitations of Dietary Records

Dietary records share some of the same limitations as recall information, due to day-to-day variability. Food intake during a single week, which provides substantially more information than a single 24-hour recall, cannot represent fully usual dietary intake. The variability of intake for specific foods is even greater than for nutrients (Salvini et al., 1988). How many and which days to record is not clear. Hackett and co-workers (1985), using 3-day records in a study among adolescents, found an unpredictable day effect and suggested that all days of the week should be sampled. Richard and Roberge (1982) observed that higher caloric intake on weekend days was related to higher alcohol consumption. Houser and Bebb (1981) evaluated variation in 3-day diary data collected from 127 participants at monthly intervals for 12 months and determined that a representative food intake must include both weekend days and weekdays, and that infrequent sampling throughout the year is more representative of usual food intake than is the same number of days collected over a short period of time. Traditionally, it has been believed that dietary records must be repeated for seasonal influence, but Kim and colleagues (1984a) did not find seasonal variation in daily food records collected for 1 year from 29 U.S. adults.

Dietary record keeping requires a literate, motivated, and cooperative population. The burden on the participant is obviously large, often imposing measuring or weighing of food, and attention to specific details of food preparation and recipe ingredients. Cooperation of the subject and validity of the record may decline in relation to the length of the recording process. Although it is difficult to document, the mere process of writing the record probably causes a change in food intake through self-observation and awareness, by omission of eating to

avoid the bothersome detail of measuring and writing, or by avoiding a normally consumed food because of embarrassment about reporting it. Furthermore, delay in recording meals and snacks obviously affects accuracy in quantification as well as memory of foods consumed.

Dietary records are expensive to process as they require the expertise of dietitians or highly trained personnel for review, decision making, and coding. Although advances in computer processing for translation to nutrient values may make this process more efficient and less prone to error, it will still remain labor-intensive. Because of extensive personnel costs and respondent burden, a frequent compromise is to decrease the number of days of recordings. Dietary records are generally not appropriate for use in case-control studies as present diet is usually not directly relevant to the etiology of most diseases once they are identified and, furthermore, the disease and its diagnosis and treatment frequently alter diet. In unusual situations, such as in the study of cancer precursor lesions where the condition being investigated could not have affected diet, diet records may be a reasonable method of dietary assessment; even then a substantial number of days are required.

APPROPRIATE APPLICATIONS OF SHORT-TERM RECALLS AND RECORDS

Appropriate Applications of 24-hour Recalls

The value of the 24-hour recall in assessing the average intake of groups is well established (Block, 1982); it can provide estimates of the average intake of large groups that are comparable to those obtained with more cumbersome techniques (Beaton et al., 1979). Twenty-four-hour recalls have been consistently used to typify food intake of large population groups such as those surveyed in HANES I and II, NFCS (The National Food Consumption Survey), and the Ten-State Survey. Because the method is essentially open-ended, it is particularly useful for assessing mean nutrient intakes among culturally different groups, and to compare nutrient intakes within groups over time. The 24-hour recall method may be particularly useful to evaluate the effectiveness of dietary intervention programs, which involves comparing the means of the intervention and control groups, as its administration should minimally alter the eating behavior of the subjects being observed. Multiple short-term recalls improve the accuracy of individual intake estimates. The effects of errors caused by intraindividual variation in dietary intake may be compensated by statistical correction procedures (see Chapter 12). The 24-hour recall may be useful for validating other dietary assessment methods among populations with limited motivation or literacy, particularly if multiple recalls are obtained for each person.

Appropriate Applications of Dietary Records

Dietary records may be used as a practical dietary assessment method with small numbers of literate, highly motivated subjects to obtain information about the frequency and consistency of eating, to characterize specific food frequencies, to

determine food quality, and to calculate nutrient data. These records may be used to monitor group compliance in dietary intervention trials; however, the possibility that compliance may be atypically good during the period of record keeping creates concerns that this information may be biased. As discussed in Chapter 6, dietary records are probably the optimal method for validating food-frequency questionnaires or dietary histories.

REFERENCES

Abraham, S., M. D. Carroll, C. M. Dresser, and C. L. Johnson (1979). *Dietary Intake Source Data: United States, 1971–74.* National Center for Health Statistics, DHEW Pub. No. (PHS) 79-1221, Hyattsville, Md.: Publie Health Serviee.

Aeheson, K. J., I. T. Campbell, O. G. Edholm, D. S. Miller, and M. J. Stock (1980). The measurement of food and energy intake in man—An evaluation of some teehniques. *Am. J. Clin. Nutr. 33,* 1147–1154.

Adelman, M. O., J. T. Dwyer, M. Woods, E. Bohn, and C. L. Otradovee (1983). Computerized dietary analysis systems: A eomparative view. *J. Am. Diet. Assoc. 83,* 421–429.

Balogh, M., H. A. Kahn, and J. H. Medalie (1971). Random repeat 24-hour dietary recalls. *Am. J. Clin. Nutr. 24,* 304–310.

Beaton, G. H., J. Milner, P. Corey, V. MeGuire, M. Cousins, E. Stewart, M. deRamos, D. Hewitt, P. V. Grambseh, N. Kassim, and J. A. Little (1979). Sourees of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *Am. J. Clin. Nutr, 32,* 2546–2549.

Beaton, G. H., J. Milner, V. MeGuire, T. E. Feather, and J. A. Little (1983). Sourees of varianee in 24-hour recall data: Implications for nutrition study design and interpretation. Carbohydrate sourees, vitamins and minerals. *Am. J. Clin. Nutr. 37,* 986–995.

Becker, B. G., B. P. Indik, and A. M. Beeuwkes (1960). *Dietary Intake Methodologies—A Review.* Ann Arbor, Mi.: U. Michigan Researeh Institute.

Bingham, S. A. and J. H. Cummings (1985). Urine nitrogen as an independent validatory measure of dietary intake: A study of nitrogen balanee in individuals eonsuming their normal diet. *Am. J. Clin. Nutr. 42,* 1276–1289.

Block, G. (1982). A review of validations of dietary assessment methods. *Am. J. Epid. 115,* 492–505.

Burke, B. S. and H. C. Stuart (1938). A method of dietary analysis. Application in researeh and pediatric praetiee. *J. Pediatrics 12,* 493–503.

Burke, B. S. (1947). The dietary history as a tool in researeh. *J. Am. Diet. Assoc. 23,* 1041–1046.

Caggiula, A. W., R. R. Wing, M. P. Norwalk, N. C. Milas, S. Lee, and H. L. Langford (1985). The measurement of sodium and potassium intake. *Am. J. Clin. Nutr. 42,* 391–398.

Campbell, V. A. and M. L. Dodds (1967). Colleeting dietary information from groups of older people. *J. Am. Diet. Assoc. 51,* 29–33.

Carroll, M. D., S. Abraham, and C. M. Dresser (1983). *Dietary Intake Source Data: United States, 1976–1980.* National Center for Health Statistics, Vital and Health Statistics, Series 11, no. 231. Washington, D.C.: Publie Health Serviee.

Carter, R. L., C. O. Sharbaugh, and C. A. Stapell (1981). Reliability and validity of the 24-hour recall: Analysis of data from a pediatrie population. *J. Am. Diet. Assoc. 79,* 542–547.

Ernst, N., M. Hjortland, J. Tillotson, and V. Grambsch (1980). The NHLBI nutrition data system. *J. Am. Diet. Assoc. 77,* 641–647.

Fanelli, M. T. and K. J. Stevenhagen (1986). Consistency of energy and nutrient intakes of older adults: 24-hour recall vs. 1-day food record. *J. Am. Diet. Assoc. 86,* 665–667.

Frank, G. C., G. S. Berenson, P. E. Schilling, and M. C. Moore (1977). Adapting the 24-hour recall for epidemiologic studies of school children. *J. Am. Diet. Assoc. 71,* 26–31.

Frank, G. C., A. T. Hollatz, L. S. Weber, and G. S. Berenson (1984). Effects of interviewer recording practices on nutrient intake—Bogalusa heart study. *J. Am. Diet. Assoc. 84,* 1432–1439.

Gersovitz, M., J. P. Madden, and H. Smiciklas-Wright (1978). Validity of the 24-hour dietary recall and 7-day record for group comparisons. *J. Am. Diet. Assoc. 73,* 48–55.

Greger, J. L. and G. M. Etnyre (1978). Validity of 24-hour recalls by adolescent females.

Guthrie, H. A. (1984). Selection and quantification of typical food portions by young adults. *J. Am. Diet. Assoc. 84,* 1440–1444.

Hackett, A. F., D. R. Appelton, A. J. Rugg-Gunn, and J. E. Eastoe (1985). Some influences on the measurement of food intake during a dietary survey of adolescents. *Hum. Nut. Applied Nutr. 39A,* 167–177.

Hankin, J. H. and R. Hunemann (1967). A short dietary method for epidemiologic studies. 1. Developing standard methods for interpreting seven-day measured food records. *J. Am. Diet. Assoc. 50,* 487–492.

Hepburn, F. N. (1982). The USDA national nutrient data bank. *Am. J. Clin. Nutr. 35,* 1297–1301.

Hoover, L. W. (1983). Computerized nutrient data bases. I. Comparisons of nutrient analysis systems. *J. Am. Diet. Assoc. 82,* 501–505.

Houser, H. B., A. I. Sorensen, A. S. Littell, and J. C. Vandervort (1969). Dietary intake of non-hospitalized persons with multiple sclerosis. 1. Food diary and coding methods. *J. Am. Diet. Assoc. 54,* 391–397.

Houser, H. B. and H. T. Bebb (1981). Individual variation in intake of nutrients by day, month, and season and relation to meal patterns: Implications for dietary methodology. In *Assessing Changing Food Consumption Patterns.* National Research Council, Committee on Food Consumption Patterns. Washington, D.C.: National Academy Press, pp. 155–179.

Jacobs, D. R., P. J. Elmer, D. Gorder, Y. Hall, and D. Moss (1985). Comparison of nutrient calculation systems. *Am. J. Epid. 121,* 580–592.

Karvetti, R. and L. Knuts (1985). Validity of the 24-hour recall. *J. Am. Diet. Assoc. 85,* 1437–1442.

Kim, W. W., J. L. Kelsay, J. T. Judd, M. W. Marshall, W. Mertz, and E. W. Prather (1984a). Evaluation of long-term dietary intakes of adults consuming self-selected diets. *Am. J. Clin. Nutr. 40* (Suppl.), 1327–1332.

Kim, W. W., W. Mertz, J. T. Judd, M. W. Marshall, J. L. Kelsay, and E. S. Prather (1984b). Effect of making duplicate food collections on nutrient intakes calculated from diet records. *Am. J. Clin. Nutr. 40* (Suppl.), 1333–1337.

Krantzler, N. J., B. J. Mullen, H. G. Schutz, L. E. Grivetti, C. A. Holden, and H. L. Meiselman (1982a). Validity of telephoned diet recalls and records for assessment of individual food intake. *Am. J. Clin. Nutr. 36,* 1234–1242.

Krantzler, N. J., B. J. Mullen, E. M. Comstock, C. A. Holden, H. G. Schutz, L. E. Grevetti, and H. L. Meiselman (1982b). Methods of food intake assessment—An annotated bibliography. *J. Nutr. Ed. 14,* 108–119.

Kruse, H. O., C. E. Palmer, W. Schmidt, and D. G. Wiehl (1940). Medical evaluation of

nutritional status. I. Methods used in a survey of high school students. *Milbrank Memorial Fund Quarterly 18,* 257–298.

Lansky, D. and K. D. Brownell (1982). Estimates of food quantity and calories: Errors in self-report among obese patients. *Am. J. Clin. Nutr. 35,* 727–732.

Liu, K., J. Stamler, A. Dyer, J. McKeever, and P. McKeever (1978). Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J. Chronic Diseases 31,* 399–418.

Madden, J. P., S. J. Goodman, and H. A. Guthrie (1976). Validity of the 24-hour recall. Analysis of data obtained from elderly subjects. *J. Am. Diet. Assn. 68,* 143–147.

Mahalko, J. R., L. K. Johnson, S. K. Gallagher, and D. B. Milne (1985). Comparison of dietary histories and seven-day food records in a nutritional assessment of older adults. *Am. J. Clin. Nutr. 42,* 542–553.

Marr, J. W. (1971). Individual dietary surveys: Purposes and methods (1971). *World Review of Nutrition and Dietetics 13,* 105–164.

McHenry, E. W. (1939). Nutrition in Toronto. *Can. Med. J. 30,* 4–13.

Mertz, W. and J. L. Kelsay (1984). Rationale and design of the Beltsville one-year dietary intake study. *Am. J. Clin. Nutr. 40* (Suppl.), 1323–1326.

Miles, C. W., B. Brooks, R. Barnes, W. Marcus, E. S. Prather, and C. E. Bodwell (1984). Calorie and protein intake and balance of men and women consuming self-selected diets. *Am. J. Clin. Nut. 40* (Suppl.), 1361–1367.

Moore, M. C., B. C. Judlin, and P. M. Kennemur (1967). Using graduated food models in taking dietary histories. *J. Am. Diet. Assoc. 51,* 447–450.

Morgan, K. J., S. R. Johnson, R. L. Rizek, R. Reese, and G. L. Stampley (1987). Collection of food intake data: An evaluation of methods. *J. Am. Diet. Assoc. 87,* 888–896.

Morgan, P. M., L. E. Demarest, W. G. Unglaub, and R. S. Hubbard (1971). Some factors for refusal to participate in nutrition surveys. *J. Nutr. Ed. 2,* 103–105.

Pao, E. M., S. J. Mickle, and M. C. Burk (1985). One-day and 3-day nutrient intakes by individuals—Nationwide food consumption survey findings, Spring, 1977. *J. Am. Diet. Assoc. 85,* 313–324.

Paul, O., M. H. Lepper, W. H. Phelan, G. W. Dupertuis, A. MacMillan, H. McKean, and H. Park (1963). A longitudinal study of coronary heart disease. *Circulation 28,* 20–31.

Pekkarinen, M., S. Kivioja, L. Jortikka (1967). A comparison of the food intake of rural families estimated by one-day recall and precise weighing methods. *Voeding 9,* 470–476.

Posner, B. M., C. L. Borman, J. L. Morgan, W. S. Borden, and J. C. Ohls (1982). The validity of a telephone-administered, 24-hour dietary recall methodology. *Am. J. Clin. Nutr. 36,* 546–553.

Prentice, A. M., A. E. Black, W. A. Coward, H. L. Davies, G. R. Goldberg, P. R. Murgatroyd, J. Ashford, M. Sawyer, and R. G. Whitehead (1986). High levels of energy expenditure in obese women. *Br. Med. J. 292,* 983–987.

Rasanen, L. (1979). Nutrition survey of Finnish rural children. VI. Methodological study comparing the 24-hour recall and the dietary history interview. *Am. J. Clin. Nutr. 32,* 2560–2567.

Richard, L. and A. G. Roberge (1982). Comparison of caloric and nutrient intake of adults during week and weekend days. *Nutr. Res. 2,* 661–668.

Salvini, S., D. J. Hunter, L. Sampson, M. J. Stampfer, G. A. Colditz, B. A. Rosner, and W. C. Willett (1989). Food-based validation of a dietary questionnaire: The effect of week-to-week variation in food consumption (abstr). *Int. J. Epidemiol,* (in press).

Schnakenberg, D. D., T. M. Hill, M. J. Kretsch, and B. W. Morris (1981). Diary-interview technique to assess food consumption patterns of individual military personnel. In *Assessing Changing Food Consumption Patterns.* National Research Council, Committee on Food Consumption Patterns. Washington, D.C.: National Academy Press, pp. 187–197.

Schucker, R. E. (1982). Alternative approaches to classic food consumption measurement methods: Telephone interviewing and market data bases. *Am. J. Clin. Nutr. 35* (Suppl.), 1306–1309.

Ten-State Nutrition Survey 1968–70 (1972). V. Dietary. DHEW Publ (HSM) 72-8133, U.S. Dept. Health, Education and Welfare, Atlanta, Ga.: Center for Disease Control.

Tillotson, J. L., D. D. Gorder, and N. Kassim (1981). Nutrition data collection in the Multiple Risk Factor Intervention Trial (MRFIT). Baseline nutrient intake of a randomized population. *J. Am. Diet. Assoc. 78,* 235–240.

van Staveren, W. A., P. Deurenberg, M. B. Katan, J. Burema, L. C. de Groot, and M. D. Hoffmans (1986). Validity of the fatty acid composition of subcutaneous fat tissue microbiopsies as an estimate of the long-term average fatty acid composition of the diet of separate individuals. *Am. J. Epid. 123,* 455–463.

Witschi, J. C., R. C. Ellison, D. D. Doane, G. L. Vorkink, W. V. Slack, and F. J. Stare (1985). Dietary sodium reduction among students: Feasibility and acceptance. *J. Am. Diet. Assoc. 85,* 816–821.

Worsley, A., K. I. Baghurst, and D. R. Leitch (1984). Social desirability, response bias and dietary inventory responses. *Hum. Nutr. Applied Nutr. 38,* 29–35.

Youland, D. M. and A. Engle (1976). Practices and problems in HANES: Dietary data methodology. *J. Am. Diet. Assoc. 68,* 22–25.

5

Food Frequency Methods

Because short-term recall and diet record methods are generally expensive, unrepresentative of usual intake, and inappropriate for assessment of past diet, investigators have sought alternative methods for measuring long-term dietary intake. Burke (1947) developed a detailed dietary history interview that attempted to assess an individual's usual diet; this included a 24-hour recall, a menu recorded for 3 days, and a checklist of foods consumed over the preceding month. This method was time consuming and expensive, and a highly skilled professional was needed for both the interview and the processing of the information. The checklist, however, was the forerunner of the more structured dietary questionnaires in use today. During the 1950s Stephanik and Trulson (1962), Heady (1961), Wiehl and Reed (1960), and Marr (1971) developed food-frequency questionnaires and evaluated their role in dietary assessment. Stephanik and Trulson found that a food-frequency questionnaire discriminated between groups of subjects defined by ethnicity, but did not consider that such a questionnaire could be useful in computing nutrient intakes. Heady (1961), using diet records collected by British bank clerks, demonstrated that the *frequencies* that foods were used correlated highly with the total *weights* of the same foods consumed over a several day period. He then designed a self-administered questionnaire for use in large populations based strictly on the frequencies that foods were eaten; unfortuntately, this questionnaire was apparently never employed for its intended use. Nichols and co-workers (1976) use a food-frequency questionnaire in the Tecumseh Heart Study population and failed to find an association between intake of fat, sugar, or starch-containing foods and level of serum cholesterol. Perhaps in part because of this failure to observe correlations with serum cholesterol, interest in food-frequency questionnaires (and nutritional epidemiology in general) waned during the early 1970s, but has greatly increased more recently. (In retrospect, correlation with serum cholesterol was an unfortunate criterion for validity as it is only moderately sensitive to changes in diet, see Chapters 1 and 6.) Multiple investigators have converged,

apparently independently, toward the use of food-frequency questionnaires as the method of dietary assessment best suited for most epidemiologic applications. During recent years substantial refinement, modification, and evaluation of food-frequency questionnaires has occurred, so that data derived from their use has become considerably more interpretable.

RATIONALE AND CONCEPTUAL BASIS

The underlying principle of the food-frequency approach is that average long-term diet, for example, intake over weeks, months, or years, is the conceptually important exposure rather than intake on a few specific days. Therefore, it may be advantageous to sacrifice precise intake measurements obtainable on 1 or a few days in exchange for more crude information relating to an extended period of time. Moreover, it is typically easier to describe one's usual frequency of consuming a food than to describe what foods were eaten at any specific meal in the past. This concept is supported by recent cognitive research (Bradburn et al., 1987) and has long been used by infectious disease epidemiologists investigating foodborne outbreaks; even when interest is focused on a specific meal, subjects find it difficult to recall their food intake at that time. Therefore, general questions are often posed as to whether a specified food is almost never eaten, or whether it is usually eaten, if available.

The basic food-frequency questionnaire consists of two components: a food list and a frequency response section for subjects to report how often each food was eaten. Questions related to further details of quantity and composition may be appended. In the following sections, considerations for the design of such questionnaires are discussed.

THE FOOD LIST

A basic decision in designing a questionnaire is whether the objective is to measure intake of a few specific foods or nutrients, or whether a comprehensive assessment of dietary intake is desired. A comprehensive assessment is generally desirable whenever possible. It is often impossible to anticipate at the beginning all the questions regarding diet that will appear important at the end of a study; a highly restricted food list may not have included an item that is, in retrospect, important. Furthermore, as discussed in Chapter 12, total food intake, represented by energy consumption, may be related to disease outcome and thus confound the effects of specific nutrients or foods. Even if total energy intake is not related to a disease outcome, adjustment for total intake may increase the accuracy of specific nutrient measurements (discussed in Chapter 11). Nevertheless, epidemiologic practice is usually a compromise between ideal and reality, and it may simply not be possible to include a comprehensive diet assessment in a particular interview or questionnaire, especially if diet is not the primary focus of the study.

A second questionnaire design issue is whether the primary objective is to rank individuals (i.e., to discriminate among subjects according to dietary intake) or to provide a measure of absolute intake. In most epidemiologic applications ranking is the primary objective, and conversion to absolute intake may even be done by post hoc statistical methods (see Chapter 12). In either circumstance, it is important to select carefully the most informative items for the food list to avoid fatigue and boredom that can impair concentration and accuracy. We have been impressed that individuals are willing to complete relatively long dietary questionnaires, probably because of strong general interest in food. Our current questionnaire, however, which includes approximately 120 food items, seems to be approaching the limit. Even in a highly motivated cohort (Willett et al., 1987), approximately 5 percent of women willing to complete a two-page, health-related mailed questionnaire did not complete an additional 61-item dietary questionnaire (unpublished data).

For a food item to be informative it must have three general characteristics. First, the food must be used reasonably often by an appreciable number of individuals. Second, the food must have a substantial content of the nutrient(s) of interest. Third, to be discriminating, the use of the food must vary from person to person. To illustrate the last point, a question about carrots would not help to rank subjects according to carotene intake if everyone ate one carrot a day. On the other hand, an item about spinach, which is often either avoided or is enjoyed and eaten frequently, may provide much more information even though it has a somewhat lower carotene content or lower average frequency of use. In theory, if the intake of two foods were very highly correlated, it would not be necessary to include them both; however, we have yet to observe a good example of this situation. These three characteristics can be summarized by calculating the contribution of each food to the between-person variance in the intake of specific nutrients; an application of this approach using stepwise regression analysis is described later. Apart from considerations of nutrient intake, foods may be included on a questionnaire on the basis of prior information, epidemiologic or otherwise, that an association might exist. For example, we have included mushrooms in a study of diet and cancer on the basis of a strong effect on stomach cancer in animals (Toth et al., 1982).

Several approaches can be used to compile a food list. The simplest is to examine published food composition tables and identify the foods that contain substantial amounts of the nutrient(s) of interest. Although rapid and simple, this strategy would lead to the inclusion of foods that have high nutrient concentration, such as brains, which are exceedingly rich in cholesterol, but are not eaten with sufficient frequency to be important.

Another approach is to start with a long list of foods that are potentially important nutrient sources, and systematically reduce this list. The original list may be derived from food composition tables, or informally with the help of an experienced dietitian. Reduction of the list can be accomplished by pilot testing of the questionnaire. The easiest method is to simply delete items that are infrequently used. This process, however, ignores the fact that foods with high between-person variation in their use are more informative than those that are

of similar average use, but used uniformly by all persons. A more sophisticated approach is to use stepwise regression analysis of pilot study data to identify the most discriminating food items (Heady, 1961).

We used stepwise regression to develop a questionnaire used in a cohort of over 100,000 women (Willett et al., 1985). The objective was to measure a fairly comprehensive list of 18 nutrients, but the space and the cost of data entry constrained the form to approximately two pages. A four-page pilot questionnaire including approximately 120 items was developed by consultation with an experienced dietitian and we conducted small-scale pilot studies to eliminate very infrequently used items. A four-page form consisting of about 100 food items was then mailed to a sample of 2000 cohort members and returned by 86 percent. For every individual we computed a total intake score for each of the 18 nutrients (computation detailed later). We then conducted a separate stepwise multiple regression anlaysis for each nutrient with the total nutrient intake as the dependent variable. In this process the computer algorithm identifies the food that explains the most between-person variance in nutrient intake as the first independent variable, the food that explains the most variance not accounted for by the first food as the second independent variable, and so on. The contribution a food makes is reflected in the change in cumulative R^2. This analysis thus identifies the foods that most discriminate between individuals, rather than those that contribute most to absolute intake. Results for animal fat, vegetable fat, cholesterol, and vitamin C are shown as examples in Table 5–1. It will be appreciated that R^2 values accumulate much more rapidly for those nutrients that are highly concentrated in a few foods (e.g., vitamin C) than those that are dispersed through many foods (e.g., protein). We were able to account for at least an R^2 of 80 percent (and often much more than 80%) for each of the 18 nutrients using only a modest number of foods. Several foods, such as meat, were important contributors to more than one nutrient. By selecting the most discriminating foods, and by collapsing several nutritionally similar items into one, we constructed a 61-item questionnaire, which was then mailed to over 120,000 cohort members (See Appendix).

Some caution should be entertained in using regression analysis to design a form. Because several hundred food variables may be included in an original long list of potential foods, some will enter as "statistically significant" predictors on the basis of chance alone. Therefore, the sample should be large, probably on the order of 1000 to 2000 rather than a few hundred. With this sample size, even unimportant contributors to cumulative R^2 may be statistically significant, but can be ignored. Even with a large sample, a few foods may occasionally make a modest contribution to R^2, but not make sense in terms of containing the nutrient being predicted or having an obvious association with a food rich in that nutrient. For example, in Table 5–1, corn entered as a modest predictor of cholesterol intake even though it contains none; this is presumably the result of correlation between the use of corn and another food, perhaps butter. It must be appreciated that this type of analysis used to identify items for inclusion on a questionnaire is in no sense a test of validity. Even though a high cumulative R^2 is obtained, the dependent variable, total nutrient score, is cal-

Table 5–1. Foods most predictive of between-person variation in animal and vegetable fat, cholesterol, and vitamin C intake

Animal fat		Vegetable fat		Cholesterol		Vitamin C	
Food	Cumulative R^2	Food	Cumulative R^2	Food	Cumulative R^2	Food	Cumulative R^2
Beef, main dish	0.31	Margarine	0.18	Eggs	0.53	Supplements	0.91
Whole milk	0.41	Cookies, ready made	0.36	Beef, main dish	0.61	Orange juice	0.93
Hard cheese	0.52	Nuts	0.49	Lean hamburger	0.68	Multi vitamins	0.94
Beef, mixed dish	0.60	Peanut butter	0.60	Corn	0.71	Fruit punch	0.96
Chicken, with skin	0.67	Chocolate candy	0.65	Chicken, with skin	0.73	Spinach, greens	0.96
Butter	0.71	Potato, corn chips	0.69	Chicken, no skin	0.76	Berries	0.96
Eggs	0.74	Pie, homemade	0.72	Beef, mixed dish	0.78	Brussel sprouts	0.96
Hot dogs	0.76	Home fried foods	0.74	Whole milk	0.80		
Ice cream	0.78	Pie, ready-made	0.76	Hard cheese	0.81		
Chicken, no skin	0.80	White bread	0.77	Liver	0.82		
Processed meats	0.81	Cake, homemade	0.78	Butter	0.83		
Hamburger, regular	0.82	French fries	0.79	Custard	0.84		
Hamburger, lean	0.83	Dark bread	0.79	Cookies, ready made	0.85		

Data are based on cumulative R^2 values in stepwise regression analysis of a semiquantitative food-frequency questionnaire pilot-tested among 1742 U.S. women.
From Willett et al., 1981.

culated using the same food items so that the cumulative R^2 would be 1.00 if all food items are entered as predictors.

A third approach to constructing a questionnaire is to use open-ended data, such as that obtained by diet records or 24-hour recalls, to identify the foods that contribute most importantly to the total absolute intake of a nutrient by the group as a whole. In this type of analysis, which has been used by Block and co-workers (1985), who employed 24-hour recall data from the NHANES study, individuals are ignored; only the pooled information is examined (by necessity when using NHANES data, as only one 24-hour collection was available per subject). Howe and colleagues (1986) have used this strategy, based on data from a comprehensive diet history, to identify a short list of foods contributing the most to intake of *N*-nitrosamines or their precursors. Similarly, Stryker (1987) used data from the 28 days of diet recording by 194 women in the Nurses Health Study validation study to compile lists of foods that contributed to the intake of 18 nutrients.

An advantage of this open-ended approach is that important contributors to nutrient intake are unlikely to be missed. Many arbitrary decisions, however, must be made regarding the collapsing of variables as the open-ended methods are typically coded in much finer detail than would be appropriate for items on a questionnaire. For example, in conducting this type of analysis using diet record data (Willett et al., 1988), over 300 codes for beef, pork, and lamb were collapsed into two categories that corresponded to our questionnaire. In addition, carefully collected open-ended methods often include mixed dishes, baked goods, and prepared foods that have been "dissected" and coded into ingredients (e.g., flour and shortening) that would not be included on a questionnaire, even though the final product (e.g., bread or cake) would be listed. Thus, the use of open-ended dietary data to identify foods for a questionnaire can require a major investment of time, and still be subject to arbitrary groupings of foods that may not correspond to the perception of persons completing the final questionnaire.

A modification of the open-ended approach would be to tally the foods from diet recalls or records collected from a sample of the study population without calculating nutrients. This might be particularly useful in situations when the size of the study or available resources would not justify a major investment in questionnaire development. An advantage of this approach is that information will be gained regarding the familiar names and descriptions of foods, which may be useful for studies among migrants or ethnic minority groups. At the same time, a tally of portion sizes of foods could be made; the use of such information is discussed later.

Byers and colleagues (1985) analyzed data to determine the relative contribution of 20 foods to absolute intake and to between-person variance expressed as cumulative R^2 (Fig. 5–1 for an example). These investigators observed that a relatively small list of foods, identified by stepwise regression, was able to explain a substantially higher percentage of between-person variance than the percentage of absolute total intake. Stryker and co-workers (1987) conducted similar analyses, based on 4-week diet records collected by 194 women, that supported the observation that a relatively short list of foods selected by stepwise

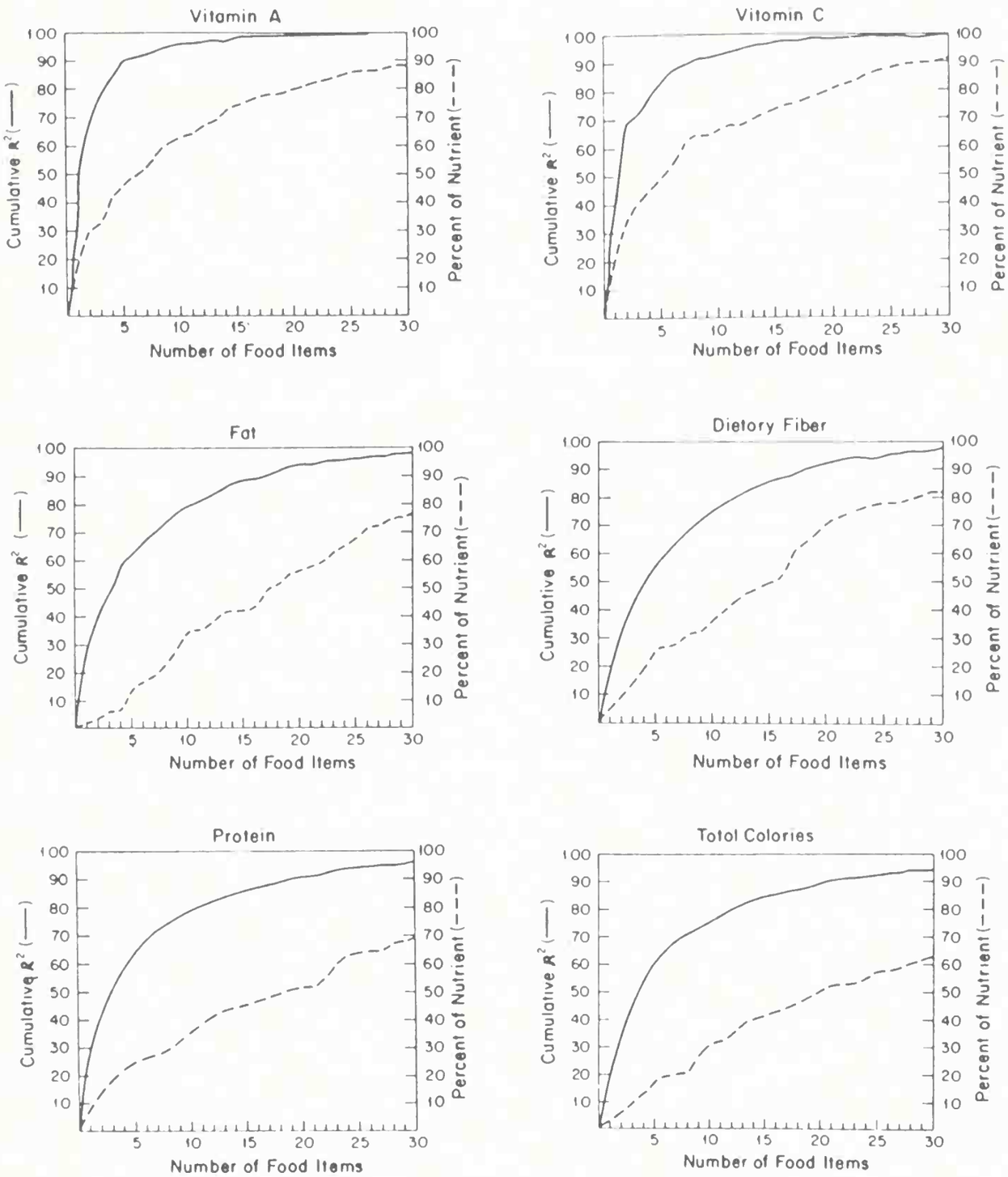


Figure 5–1. Cumulative R^2 and percentage of nutrients accounted for according to the number of food items included in the nutrient index. (From Byers et al., 1985; reproduced with permission.)

regression could account for a high portion of between-person variation. Some foods, such as meat, which was eaten regularly by nearly all persons and was the most important single contributor to absolute intake of calories and total fat, contributed less impressively to between-person variance in these nutrients than did foods such as cake and cookies, which were eaten frequently by some but avoided by other participants. The contribution of supplements to between-person variance was particularly striking, although not unexpected as their use is an all-or-nothing phenomenon. These data suggest that the method of identifying foods simply on the basis of their contribution to absolute intake may not lead to the optimal questionnaire when the number of items is constrained by cost, space, or time. Furthermore, foods or groups of foods that are most dis-

criminating among persons should probably be asked about with greater detail than those that are less discriminating.

Assembling a selected list of foods into a clear and unambiguous questionnaire is critical. To some extent this is a process of trial and error, but the following considerations may be helpful. The organization and structure of a food list is important because one item can change the interpretation of another. For this reason, related items should be clustered together, such as by traditional food groups. For closely related foods, more specific items should precede general items (e.g., "low calorie salad dressing" should appear first, then "other salad dressing"). It is tempting to maximize the comprehensiveness of a questionnaire while maintaining brevity by combining or collapsing several foods into single questions. Our experience suggests that this should be avoided; multiple simple, clear questions are preferable to a single longer, complex question. For example, on our original questionnaire we asked how often "peaches, apricots or plums" were eaten over the past year. If more than one of these fruits are eaten it is a very challenging task for respondents to consider each separately, integrate information over a year, and then describe a summary frequency. It would probably be better to ask three separate questions or drop one or two of the fruits entirely. In a related example, we found that the deletion of a clarifying item can alter response to another item. In the stepwise regression analysis described previously, we found that lettuce did not contribute importantly to variance for any nutrient and so we dropped this item. We had, however, included another item "Spinach and other greens" that was an important predictor of carotene intake. With the lettuce item deleted, we discovered that reporting of "spinach and other greens" increased, apparently because some may have interpreted this to include lettuce. Although we thought we may have been gaining information on possible use of kale, collard, and mustard greens with our more comprehensive question, we probably gained statistical noise instead; it would have been better to simply ask about "spinach." Arriving at the optimal balance between splitting foods into many detailed separate questions or collapsing (combining) items into a few broad questions is extremely important. Because it is critical that the person completing the questionnaire has the same mental picture of the foods as intended by the investigator, simplicity and clarity are paramount. Cognitive research indicates that subjects often decompose a question into several steps when they are unable to retrieve and count separate incidents, such as the repeated use of a food (Bradburn et al., 1987). Thus, much may be learned from pilot studies that include informal or structured interviewing of participants as to their interpretation of questionnaire items.

An alternative to designing a questionnaire de novo is to use or modify an existing questionnaire. This will be particularly attractive when time (often years) and resources are not avilable for a major investment in questionnaire development. There appears to be general willingness on the part of epidemiologists to share their questionnaires for legitimate scientific purposes. If the hypotheses are adequately addressed by an existing questionnaire and the form is reasonably appropriate for the cultural background of the study population, this may be the best approach. When the number of questions is severely con-

strained and the focus is on one or a few nutrients only, it may be reasonable to select a list of foods identified by other investigators as important predictors if the general cultural background of the study population is similar.

FREQUENCY RESPONSE SECTION

For most epidemiologic purposes, dietary intake over a number of years is the exposure of conceptual interest. Because diets tend to be reasonably correlated from year to year (see Chapter 8), we and most others have asked subjects to describe their frequency of using foods in reference to the preceding year. This provides a full cycle of seasons so that, in theory, the responses should be independent of the time of year. For other purposes, the time frame could be in reference to a period 5 years previously (such as in a case-control study of colon cancer), the first 2 months of pregnancy (in a study of congenital malformations), or the preceding month (in a study of plasma HDL–cholesterol). The relevant reference period is a function of both the physiology or the pathophysiology of the outcome being studied as well as the metabolism of the dietary factor under investigation. In the case of congenital malformations, the occurrence of the critical event, such as the closing of the neural tube, can be identified with precision to within several days. Because most nutrients (e.g., vitamin A) have long half-lives in the body, however, the nutritional status at the time of neural tube closure will depend on intake over the preceding weeks or months.

Response formats to frequency questions may appear to be a simple issue, but opportunities exist for serious pitfalls. Most investigators have provided a multiple-choice response format, with the number of options usually ranging from five to ten. If the same options are to be used for all food questions, however, five choices are likely to be too few and will usually result in serious loss of information. This can be appreciated by examining the frequency distributions for selected items in our pilot study of 1742 women (Table 5–2). For some items (such as butter), most of the information on between-person variation is captured at the high end of the scale, but for other items (e.g., liver), most of the information is at the low end. Broadening the response categories to reduce their number would decrease the discrimination capacity of questions. For example, Gray and colleagues (1984) compared a food frequency questionnaire having a small number of response options with a detailed dietary interview and observed a poor correlation. They attributed this to the limited response options, as for many items most subjects fell into a single category that encompassed a possible fourfold variation of intake. In an early questionnaire, Stephanik and Trulson (1962) perceptively created ten categories so that the range of intake within a category was no more than twofold:

never
once a month or less
2–3 times per month
once per week
2–4 times per week

Table 5–2. Frequency distributions of responses to selected foods

Food	Percent of Women								
	Almost never	1–3/mo	1/wk	2–4/wk	5–6/wk	1/day	2–3/day	4–6/day	6+day
Skim/low fat milk	45	6	4	9	5	19	12	0	0
Whole milk	51	8	7	6	2	14	6	1	0
Yogurt	61	17	9	6	1	2	0	1	0
Ice cream	22	29	23	19	3	3	0	0	0
Butter	56	4	6	7	3	8	12	1	0
Apples or pears	14	18	19	27	4	13	4	0	0
Fish	9	34	39	15	3	0	0	0	0
Cabbage, cauliflower	15	36	34	12	1	1	0	0	0
Liver	33	64	3	0	0	0	0	0	0
Coffee	23	2	1	3	2	11	33	17	8
Low-calorie carbonated beverage	49	9	7	11	5	10	7	1	0

Data are based on a pilot study conducted among 1742 women. From Willett et al., 1981.

See Appendix 5–1 for data on additional foods.

5–7 times per week
over one, under two times a day
2–3 times per day
4–6 times per day
over 6 times per day

This produces a scale with greater detail at the high frequency end, which is appropriate as foods consumed less than once per week make relatively little contribution to nutrient intake. We have modified these categories slightly (see appendix 5–1), collapsing two low-frequency responses into one, thus more efficiently using contemporary data-processing facilities (because nine possible responses plus "blank" use a single column of data storage).

One approach is to use an open-ended format and provide subjects the option of answering in terms of frequency per day, week, or month (Block et al., 1986). In theory, an open-ended frequency response format might provide for some enhanced precision in reporting as the frequency of use is truly a continuous rather than a categorical variable. It is unlikely, however, that the overall increment in precision is large as the estimation of the frequency a food is used is inherently an approximation. Moreover, tendencies to overestimate or underestimate will in part be compensated for by adjusting for total food intake (see Chapter 11). Nevertheless, formal investigation of possible refinements in response formats is warranted.

OPTIONS FOR PORTION SIZE INFORMATION

Whether or not to collect additional data on portion sizes has been a controversial topic; relevant data are only now becoming available. Several options exist. The first is to collect no additional information on portion sizes, that is, to use a simple frequency questionnaire. A second possibility is to specify a portion size as part of the question on frequency, that is, to ask how often a glass of milk is consumed rather than only how often milk is consumed. This has been termed a semiquantitative food frequency questionnaire. For foods that come in natural units (such as a slice of bread, one egg, a cup of coffee) this additional specification can add clarity to the question. For example, if we just ask about "milk," a thoughtful subject will not know if we are asking about the milk added to coffee and breakfast cereal, or whether three glasses consumed at one meal should count as one or as three. Because the specification of unit size does not require an additional question and should provide additional information and clarity, there is little reason not to pose questions in this way. The issue is less clear for foods that do not come in natural or typical units, such as meat or rice. It is possible to specify a typical portion (e.g., 4 to 6 ounces for a serving of meat or half a cup of rice); if a subject's usual portion is twice that amount, they would be expected to double their reported frequency of use. We have informally evaluated whether participants in the Nurses Health Study Cohort responded in this way by providing them with a hypothetical eating pattern and asking them to complete questions describing the diet. Although items about foods with natural

units were consistently interpreted correctly, the portion size specification for foods without natural units, such as meat, was sometimes ignored. Whether or not it is advantageous to include a portion size specifically for these foods deserves further investigation (discussed later).

A third alternative is to include an additional item for each food to describe the usual portion size. This may involve describing a typical portion size in words and asking subjects to describe their usual portion as a multiple of the specified portion (Block et al., 1986), using a realistic food model or a simple shape as a unit of reference (Morgan et al., 1978), or providing pictures of different portion sizes as a multiple-choice question (Hankin et al., 1983). The descriptions of usual serving sizes can also be left as completely open-ended items, although this will be expensive to code and process.

To provide useful information on serving sizes, subjects must be able to conceptualize the unit clearly and relate this to their own habits. Guthrie (1984) has observed, however, that individuals are generally unable to accurately describe their portion sizes. When subjects were asked, immediately after a meal, to describe the portions of foods they had just consumed, for most of the foods evaluated, fewer than half of the participants were accurate to within 25 percent. Furthermore, substantial within-person variation exists in portion sizes for most foods (Hunter et al., 1988). For example, beef is eaten in a variety of forms and circumstances in which the portion size may vary substantially.

In Table 5–3 the within-person and between-person coefficients of variation for portion sizes of several commonly eaten foods are provided. Among a total of 66 foods examined, the average ratio of the within-person to between-person variance was approximately four (Hunter et al., 1988). This large within-person variation raises doubt about the concept of "usual" portion size. Indeed, in some situations we would not really be interested in the mode serving size for an individual if it was substantially smaller than another commonly used serving size. For example, a common serving size reported for yogurt is one tablespoon. Even if yogurt is consumed as a tablespoon serving size at, say, three times the frequency as a cup, we would still be more interested in learning about the frequency that the cup serving was used as it will provide a greater source of nutrients.

Samet and colleagues (1984) have addressed the contribution of portion-size questions to the ranking of individuals. They calculated intake of vitamin A in two ways; the first was based on simple frequency questions alone and the second by weighting frequency responses by "usual" portion sizes obtained for these same foods during an in-person interview that employed food models. The correlation between the methods was 0.86 for controls and 0.91 for cases, indicating that the portion-size questions provided little additional information. Furthermore, vitamin A intake calculated with and without portion sizes was similarly related to a reduced risk of lung cancer (Humble et al., 1987).

We have conducted similar analyses using data collected as part of a case-control study of coronary heart disease (unpublished data). Participants completed a mailed self-administered food-frequency questionnaire that specified units for foods commonly consumed in "natural" units but did not specify portions for items such as meat, cooked vegetables, rice, and mashed potatoes. After

Table 5–3. Within-person and between-person coefficients of variation in portion sizes (CV_w and CV_b) of selected commonly consumed foods among 194 women

Food item	Portion	Mean	Median	CV_w (%)	CV_b (%)	Variance ratio[a]	Variance ratio (ln)[b]
Dairy Foods							
Yogurt	cup	0.5	0.5	39.8	61.9	0.4	0.8
Ice cream	cup	0.6	0.5	43.1	23.1	3.5	3.7
Cottage cheese	cup	0.5	0.4	75.6	45.4	2.8	3.1
Hard cheese	oz	1.0	1.0	73.3	24.1	9.3	9.5
Butter	tsp	1.9	1.2	78.3	33.1	5.6	5.3
Meats							
Processed meat	oz	2.0	1.8	62.0	23.0	7.2	4.9
Hamburgers	oz	2.9	2.6	54.7	33.0	2.7	3.1
Chicken	100 g	0.9	0.8	53.9	24.5	4.9	5.4
Red meat	100 g	1.0	0.9	60.2	22.0	7.5	11.3
Vegetables							
Broccoli	cup	0.4	0.5	54.7	25.0	4.8	4.3
Cabbage/ cauliflower	cup	0.6	0.5	63.0	25.3	6.2	7.1
Carrots	cup/one	0.5	0.4	75.3	33.6	5.0	6.0
Corn	cup/one	1.0	0.5	85.8	56.1	2.3	3.3
Spinach	cup	0.6	0.5	79.0	28.2	7.9	2.6
Peas	cup	0.4	0.5	44.5	25.3	3.1	4.5
Potatoes mashed/baked	cup	0.7	0.8	36.2	19.6	3.4	5.7
Lettuce	cup	0.9	0.7	61.6	36.0	2.9	3.0
Tomatoes	raw, one	0.4	0.4	63.4	28.4	5.0	5.1
Fruit							
Peaches	cup	0.5	0.5	45.9	25.0	3.4	2.0
Raisins	oz	0.4	0.3	95.0	48.5	3.8	1.9
Canteloupe	cup	0.7	0.7	49.3	31.0	2.5	2.5
Fish							
Tuna fish, canned	100 g	0.7	0.6	37.2	27.2	1.9	1.6
Dark fish	100 g	1.1	1.0	47.4	32.1	2.2	2.5
White fish	100 g	1.3	1.1	37.9	21.7	3.1	4.4
Breads, Cereals, Starches							
Cake	oz	2.3	1.9	61.2	26.0	5.5	6.1
Pizza	⅛	2.1	2.0	55.1	11.7	22.4	93.2
Pasta/spaghetti noodles	cup	1.0	1.0	54.5	24.9	4.8	4.8
Cold cereal	cup	0.8	0.8	43.3	52.2	0.7	0.5
Cooked cereal	cup	0.7	0.8	22.9	25.6	0.8	0.8
Potato/corn chips	oz	1.0	0.8	57.4	36.6	2.5	1.7
Popcorn	cup	2.7	2.0	81.4	58.1	2.0	1.9
Crackers	one	4.4	4.0	100.8	47.9	4.4	3.5
Pancakes/waffles	one	2.4	2.0	40.4	41.2	1.0	0.4
Bagel/muffins	one	1.1	1.0	40.8	23.4	3.1	3.0

[a]Variance ratio = *Within-person variance*/Between-person variance = s_w^2/s_b^2

[b]Based on $\log_e$ transformed data.

From Hunter et al., 1988.

completing the questionnaire, subjects were interviewed by a dietitian who used realistic food models to estimate usual portion sizes. Nutrient intakes were then calculated in two ways: first, assuming common portion sizes for all subjects and, second, using the portion size data obtained at the interview. For each nutrient examined, the correlation between these methods of computation was greater than 0.90 (Table 5–4). Ninety-eight male control subjects in this study also collected a 1-week diet record after completing the questionnaire. Nutrients calculated with and without the additional portion-size data correlated equally well with those estimated by the diet record (Hernandez-Avila et al., 1988). Although it remains possible that a more detailed assessment of usual portion sizes would slightly improve measurements of nutrient intake, it is apparent that the data collected (at considerable expense) during the personal interview added little to the assessment of nutrient intake. In a study among 37 postmenopausal women, Cummings and co-workers (1987) found that calcium intake calculated from a questionnaire, which included ratings of portion sizes as small, medium, and large, correlated only somewhat and nonsignificantly more with calculations from a diet record ($r = 0.76$) than did the questionnaire without added information on serving sizes ($r = 0.64$). A questionnaire completed by the same women that asked for portion sizes in ounces or cups performed less; the correlation with diet records was 0.49. Although Hankin and colleagues have suggested that additional questions on portion sizes may be important, they found that the use of data obtained using color photographs of different portion sizes only slightly increased correlations between intake of foods assessed by a ques-

Table 5–4. Spearman correlation coefficients comparing semiquantitative food-frequency questionnaire nutrient intakes with and without data on portion sizes

Nutrient	Cases (n = 115)	Controls (n = 114)	Cases + controls (n = 229)
Protein	0.94	0.94	0.94
Animal fat	0.94	0.93	0.93
Vegetable fat	0.99	0.99	0.99
Saturated fat	0.97	0.96	0.97
Oleic acid	0.95	0.96	0.96
Linoleic acid	0.98	0.98	0.98
Crude fiber	0.93	0.94	0.94
Dietary fiber	0.93	0.95	0.94
Preformed vitamin A	0.99	0.95	0.99
Carotene	0.90	0.91	0.91
Cholesterol	0.96	0.96	0.96
Vitamin C	0.97	0.98	0.97
Vitamin B_6	0.93	0.95	0.94
Vitamin E	0.98	0.99	0.99
Sucrose	0.99	0.99	0.99
Total carbohydrate	0.97	0.98	0.98
Total calories	0.97	0.97	0.97

Calculated assuming common portion sizes for all subjects and using data for individual portion sizes collected during an interview.

Unpublished data derived from a case-control study of myocardial infarction.

tionnaire completed at an interview and by a 1-week diet record (Hankin et al., 1975). For 30 food items examined, the average Spearman correlation only increased from 0.55 to 0.59.

As pointed out by Samet and colleagues (1984) and Pickle and Hartman (1985), for most foods, portion sizes vary less among individuals than do frequencies of use. Because most of the variation in intake of any food is explained by frequency of use, it is not surprising that portion-size data are relatively unimportant. Furthermore, these investigators found that portion sizes were positively correlated with frequency of use, meaning that some of the information on portion size was already accounted for by frequency of use. The finding by Hunter and co-workers (1988) that, for most foods, portion sizes vary considerably more within a person over time than between persons of similar age and sex, means that the concept of usual portion size is inherently complex and unlikely to be reported precisely.

Although the potential contribution of additional questions on portion size deserves further investigation, available data suggest that such questions do not add substantially to the assessment of dietary intake. This has important implications for study design as the cost of data collection by mail or telephone is far less than the cost of personal interviews, which are necessary if food models are to be used for assessing portion sizes. It should be noted that the relative importance of portion sizes are, in part, culturally based. It is possible to conceive of other situations: for instance, a culture in which meat is usually eaten once a week, but the amount eaten on that day varies substantially between persons; in this case, a question focused on the amount of meat eaten could be quite informative.

COMPUTATION OF NUTRIENT INTAKES

Although the data on frequency of food use obtained from a questionnaire is intrinsically useful, most investigators also want to examine the relationships of nutrient intakes with health outcomes. A nutrient database and analysis program must be compiled to calculate these intakes. In constructing the database, a value for each nutrient being computed must be assigned to each food. If portion sizes have been specified on the questionnaire, the nutrient values relate to that portion size. If no portion sizes have been specified, the nutrient content for a typical or average portion size should be used. Sources of nutrient composition data and their limitations are discussed in Chapter 2; because most published databases are incomplete, it is often necessary to use more than one source for less common nutrients when compiling the database for a specific study. If open-ended questions are included, such as for types and brands of breakfast cereal or multiple vitamins, it is necessary to obtain specific data for each possible response.

Total intake of a nutrient can be calculated as the sum of the products of the frequency weight and the nutrient content for each food, i.e., Σ (frequency weight $\times$ nutrient content). For frequency weights, it is simplest to assign a

	FOODS AND AMOUNTS	Never or less than once per mo	1-3 per mo	1 per wk	2-4 per wk	5-6 per wk	1 per day	2-3 per day	4-5 per day	6+ per day
A	Eggs (1)	○	○	Ⓦ	○	○	●	○	○	○
B	Whole milk (8 oz glass)	○	○	Ⓦ	○	○	Ⓓ	●	○	○
C	Ice cream (½ cup)	○	○	Ⓦ	○	●	Ⓓ	○	○	○

Figure 5–2. Example of calculation of daily cholesterol intake. From a food composition table the cholesterol contents are: 1 egg = 274 mg, 1 glass of milk = 33 mg, ½ cup of ice cream = 29.5 mg. Thus, the average daily cholesterol intake for the person completing this abbreviated questionnaire would be: 274 mg × 1 + 33 mg × 2.5 + 29.5 mg × 0.8 = 380.1 mg/day. (From Sampson, 1985, reproduced with permission.)

weight of 1.0 to once a day, and proportional weights to the other responses, that is, "2–3 times a day" = 2.5. The calculation of intake based on a section of a semiquantitative food-frequency questionnaire is shown in Figure 5–2. If separate portion-size questions were asked, the product for each food would also be multiplied by a weight proportional to the usual serving size.

In more complex questionnaires, the nutrient content of some foods may be modified by responses to other questions. For example, the composition of the cold breakfast cereal item could be modified by the brand and type of cereal, or the values of margarine might be modified by whether stick or tub margarines are usually used. Similarly, the nutrient values for baked goods made at home may be modified by the type of fat used (e.g., margarine vs. butter), which, in turn, can be modified by the type of margarine.

Because opened-ended questions relating to specific types or brands of foods, such as for cold breakfast cereal or margarine, require an additional coding step, they add appreciably to the cost of processing a questionnaire. We, therefore, examined the increment in validity provided by open-ended questions for type of breakfast cereal, multiple vitamins, and cooking oil. Nutrient intakes computed with and without them (assuming a common value for all breakfast cereals, multiple vitamins, and cooking oil) were compared with intakes measured by diet records (Willett et al., 1988). In the case of breakfast cereals and cooking oil, correlations for each nutrient examined were essentially unchanged; the type of multiple vitamin did contribute to the estimation of vitamins and iron (e.g., the correlation for iron was 0.55 with the brand and type of multiple vitamin and 0.42 when a common type of vitamin was assumed for all subjects). The added costs of open-ended questions related to details of foods and supplements thus appear to provide only a small increment to the estimation of nutrient intake if the basic questionnaire is reasonably complete. This issue, however, deserves further examination as it might not apply to other populations and for all nutrients.

USE OF EMPIRICAL PREDICTOR SCORES

An alternative to the calculation of nutrient intakes from food-frequency questionnaires using values from published food composition tables is to employ empirically derived weights for each food item. The rationale for such a procedure is that the importance of a food item should reflect not only the nutrient content of the food, but also the validity of the responses to that particular item. Calculation of empirical weights require that an external measurement of intake be available for a group of subjects who completed the questionnaire; this independent measurement could be based on either diet-records or a biochemical parameter, such as plasma beta-carotene. Weights for the questionnaire food items can be obtained by fitting a multiple regression model with the external measurement as the dependent variable and the food items as the independent variables. The coefficients for the food items would serve as weights that could then be applied to questionnaire responses in other studies.

This empirical approach has the advantage that it avoids many assumptions regarding portion sizes, nutrient composition, and bioavailability. Furthermore, questions that are unclear and thus improperly answered tend to receive low weights, thereby reducing error as persons who might have been erroneously assigned extreme values will be given a more average value. The use of empirical weights is limited largely by practical considerations. For their computation, the regression coefficients for individual foods must be estimated with considerable precision; it is of minimal use to simply know whether they are significantly different from zero. When many foods contribute to intake of a nutrient, the coefficient for any single food is relatively small and a number of variables are likely to be statistically significant on the basis of chance alone. For these reasons the generation of reasonably precise empirical predictor weights requires a large sample, probably many hundreds or several thousand subjects. Unless a large sample is used, weights corresponding to values from food composition tables will probably be more accurate than the empirical weights. In an exploratory analysis, we found that empirical weights provided a better prediction of plasma vitamin E (alpha-tocopherol) levels when tested in an independent dataset than did weights from a food composition table (unpublished data). The appropriate role for empirical weights is not yet clear and deserves examination in other settings.

OVERALL QUESTIONNAIRE DESIGN AND ADMINISTRATION

Basic principles of survey research (Rossi and Wright, 1985; Bradburn et al., 1987; Worsley, 1981) are likely to apply to the collection of food frequency data; however, this method of dietary assessment does not appear to have received formal investigation by this field. Although a detailed discussion of these principles is beyond the scope of this book, Babor and colleagues (1987) have provided an outline of sources of bias and solutions (Table 5–5).

Table 5–5. Procedures for minimizing response bias and enhancing validity

Source of bias	Solution
Ambiguous role requirements or task definition	Guarantee anonymity/confidentiality
	Give clear instructions/examples
	Emphasize scientific role
Specificity of desired responses	Question wording
	Multiple questions
	Probe questions
Interviewer bias and unreliability	Interviewer training
	Standardized protocols
	Interview techniques
Forgetting, telescoping	Increase question length
	Better instructions
	Aided recall
	Memory aids
	Bounded recall
Response distortion	Alert subject to problem
	Recognition questions
	Diaries/calenders/time-line
Motivation	Commitment agreement
	Clear instructions
	Bounded recall
	Prompts/encouragement/reinforcement
	Session feedback
	Bogus pipeline

From Barbor et al., 1987.

The accurate collection of dietary data requires a motivated subject; thus, emphasizing the scientific importance of the information and assuring confidentiality is basic. Clear instructions are essential, which can usually be enhanced by the use of relevant examples. As already discussed, clarity of questions is paramount, which can often be improved by feedback from subjects during questionnaire development.

If a questionnaire is to be used in an interview format, standardized procedures are essential. Interestingly, Babor and co-workers (1987) suggest that longer questions may be helpful in an interview as this may provide subjects with a longer time to think about their response. The use of memory aides and prompts, perhaps related to well-known events, has been used for many types of surveys, but their utility for the recall of dietary data remains to be documented. Among the recall aides, Babor and colleagues note that fixed response choices, as we have employed for frequency-of-use, can reduce memory errors.

A number of the other suggestions offered by Babor and colleagues have not been used or evaluated in the collection of food-frequency data. One interesting approach, which they have called *bogus pipeline,* is to ask subjects to provide information under conditions where they are led to believe that objective, external validation of their responses, such as by blood test, will also be available to the investigator. This technique has been proven to enhance the validity of reporting of alcohol intake. Although conveying "bogus" information to study participants would trouble most epidemiologists, circumstances may actually

exist in dietary studies where blood or other tissues are being collected, frequently to assess a limited number of nutrients; explicitly linking these sources of data when describing the study might be useful. Although most of the general principles of survey research are already incorporated into standard epidemiologic practice, some of the methods developed to enhance recall of past information deserve consideration and evaluation in the context of food-frequency questionnaires.

SUMMARY

Food-frequency questionnaires have become the primary method for measuring dietary intake in epidemiologic studies. Such questionnaires are directed to the dietary exposure of conceptual interest in most applications, which is average intake over an extended period of time. Food-frequency questionnaires are extremely practical in epidemiologic applications as they are easy for subjects to complete, often as a self-administered form. Processing is readily computerized and inexpensive, so that even prospective studies involving tens of thousands of subjects are feasible. In constructing food-frequency questionnaires, careful attention must be given to the choice of foods and the format of the frequency response section. The actual performance of such questionnaires in terms of reproducibility and validity is considered in the next chapter.

APPENDIX: 1980 NURSES' HEALTH STUDY DIETARY QUESTIONNAIRE WITH FREQUENCY DISTRIBUTION OF RESPONSES (%)

For each food listed, check the box indicating how often, **on average**, you have used the amount specified **during the past year**. If your intake of a food item has greatly increased or decreased during the past 10 years, indicate this in the last 2 columns.

	Average use last year								
FOOD AND AMOUNTS	6+ per day	4-6 per day	2-3 per day	1 per day	5-6 per week	2-4 per week	1 per week	1-3 per month	Almost Never
Dairy Foods									
Skim or low fat milk (8 oz. glasses)	0	1	12	20	5	12	6	5	39
Whole milk (8 oz. glasses)	0	0	4	10	2	8	7	7	62
Yoghurt, (1 c.)	0	0	0	2	1	8	9	20	61
Ice cream (½ c.)	0	0	0	3	3	18	22	32	22
Cottage cheese (½ c.)	0	0	1	3	3	18	17	29	28
Hard cheese, plain or as part of a dish (slice or servings)	0	0	3	12	13	36	20	11	5
Margarine (pats added to food or bread)	2	4	28	23	9	11	4	3	17
Butter (pats added to food or bread)	1	2	10	10	4	7	4	5	58
Fruits									
Fresh apples or pears (1)	0	2	3	15	6	27	19	20	11
Oranges (1)	0	0	2	11	4	22	19	23	19
Orange or grapefruit juice (small glass)	4	3	4	35	9	19	10	10	13
Peaches, apricots or plums (fresh, ½ c. canned, or dried)	0	1	1	3	2	14	19	32	30

FOOD AND AMOUNTS	Average use last year								
	6+ per day	4-6 per day	2-3 per day	1 per day	5-6 per week	2-4 per week	1 per week	1-3 per month	Almost Never
Bananas (1)	0	0	0	6	4	23	25	27	15
Other fruits (fresh, or ½-c. canned)	0	0	2	9	6	24	23	23	14
Vegetables String beans (½-c.)	0	0	0	2	3	30	46	16	3
Broccoli (½-c.)		0	0	1	1	16	42	29	11
Cabbage, cauliflower, brussels sprouts (½-c.)	0	0	0	1	1	12	34	37	15
Carrots (whole or ½-c. cooked)	0	0	0	3	4	21	39	26	7
Corn (ear or ½-c.)	0	0	0	0	1	13	38	32	15
Spinach or other greens (½-c.)	0	0	2	10	8	21	30	20	10
Peas or lima beans (½-c. fresh, frozen or canned)	0	0	0	1	1	15	39	29	15
Yellow (winter) squash (½-c.)	0	0	0	0	0	4	14	34	48
Sweet potatoes (½-c.)			0	0	0	1	4	29	67
Beans or lentils, dried (½-c.)	0	0	0	0	0	3	12	36	48
Tomatoes (1) or tomato juice (4 oz.)	0	0	1	11	12	33	24	15	4
Meats Chicken, without skin (6-8 oz.)	0	0	0	1	2	19	36	16	26
Chicken, with skin (6-8 oz.)		0	0	0	1	11	35	16	37
Hamburgers (1)	0	0	0	0	1	22	52	19	5
Hot dogs (1)	0	0	0	0	0	5	27	37	30
Processed meats (sausage, salami, bologna, etc.) (piece or slice)	0	0	0	2	3	15	20	27	32
Bacon (2 slice servings)	0	0	0	1	1	9	24	33	31
Beef, pork or lamb as a sandwich or mixed dish (stew, casserole, lasagne, etc.)	0	0	0	2	3	25	34	26	9
Beef, pork or lamb as a main dish (steak, roast, ham, etc. 6-8 oz.)	0	0	1	5	10	41	31	10	3
Fish (6-8 oz.)	0	0	0	1	2	15	39	34	9
Eggs (1)	0	0	2	8	9	48	20	9	5
Sweets, Baked Goods, Cereals Chocolate (1 oz.)	0	0	1	4	3	12	18	28	34
Candy without chocolate (1 oz.)	0	0	0	2	1	6	13	26	51
Pie, home made (slice)			0	0	0	1	8	40	50
Pie, ready made (slice)		0	0	0	0	1	4	21	74
Cake, (slice)	0	0	0	1	1	7	17	45	30
Cookies (1)	0	1	5	5	9	20	16	24	20
Cold breakfast cereal (½-c.)	0	0	0	10	6	21	13	14	35
White bread (slice)	0	2	17	15	9	17	7	6	27
Dark or whole grain bread (slice)	0	1	14	17	10	20	9	12	17
Miscellaneous Peanut butter (tbsps)	0	0	1	4	4	13	15	23	40
Potato or corn chips (small bag or 1 oz.)	0	0	0	1	2	8	17	28	44
French fried potatoes (4 oz.)	0	0	0	0	0	4	19	36	41
Nuts (1 oz.)	0	0	1	2	2	9	15	36	34

FOOD AND AMOUNTS	Average use last year								
	6+ per day	4-6 per day	2-3 per day	1 per day	5-6 per week	2-4 per week	1 per week	1-3 per month	Almost Never
Potatoes, mashed (½-c.) or baked (1)	0	0	0	4	7	35	26	18	9
Rice or pasta (½-c.)	0	0	0	1	2	27	38	23	9
Coffee, not decaffinated (cups)	8	17	33	11	2	3	1	2	23
Tea (cups)	1	4	15	16	4	10	8	11	30
Beer (bottles or cans)	0	0	1	1	1	4	5	9	78
Wine (glasses)	0	0	2	6	3	11	10	23	44
Liquor - whiskey, gin, etc. (drinks)	0	0	2	4	2	8	8	18	56
Coca Cola, Pepsi, other cola (glasses)	0	0	2	5	3	9	9	14	57
Low calorie carbonated drink (glasses)	0	1	7	10	5	11	6	9	50
Other carbonated beverage (root beer, ginger ale, 7-Up, etc.) (glasses)	0	0	1	2	1	6	9	20	62
Fruit-flavored punch or non-carbonated beverage (glasses)	0	0	1	2	2	6	6	12	71
Home-fried food, any type (servings)	0	0	0	2	3	15	18	23	38
Artificial sweetner (packet, tablets, etc.)	2	4	9	6	2	4	2	4	68

How often do you eat liver (3-4 oz servings)? [3] 1 per week [12] 2-3 per month [52] 1 per month or less [33] never

What do you do with the visible fat on your meat? [4] eat most of it [20] eat some of it [76] eat as little as possible

What kind of fat do you usually use for baking? [4] lard or butter [27] vegetable oil [29] vegetable shortening [27] margarine

What kind of fat do you usually use for frying? [3] lard or butter [61] vegetable oil [15] vegetable shortening [13] margarine

Do you use a microwave oven? [27] Yes [73] No If yes, for how many years? [ ]

Are you currently on a special diet? [17] Yes [83] No If yes, for [ ] years type of diet _____

In what form do you usually use your margarine? [66] Stick form [34] Tub form

Do you currently take any of the following vitamins?	YES	NO
Multiple vitamins	[34]	[86]
Vitamin A	[4]	[96]
Vitamin C	[19]	[81]
Vitamin E	[13]	[87]

REFERENCES

Babor, T. F., R. S. Stephens, and G. A. Marlatt (1987). Verbal report methods in clinical research on alcoholism: Response bias and its minimization. *J. Studies on Alcohol 48*, 410–424.

Block, G., A. M. Hartman, C. Dresser, et al. (1986). A data-based approach to diet questionnaire design and testing. *Am. J. Epidemiol. 124*, 453–469.

Block, G., C. M. Dresser, A. M. Hartman, and M. D. Carroll (1985). Nutrient sources in the American diet: Quantitative data from the NHANES II survey. I. Vitamins and minerals. *Am. J. Epidemiol. 122,* 13–26.

Burke, B. S. (1947). The dietary history as a tool in research. *J. Am. Diet. Assoc. 23,* 1041–1046.

Bradburn, N. M., L. J. Rips, and S. K. Shevell (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science 236,* 157–161.

Byers, T., J. Marshall, R. Fiedler, M. Zielezny, and S. Graham (1985). Assessing nutrient intake with an abbreviated dietary interview. *Am. J. Epidemiol. 122,* 41–50.

Cummings, S. R., G. Block, K. McHenry, and R. B. Baron (1987). Evaluation of two food frequency methods in measuring dietary calcium intake. *Am. J. Epidemiol. 126,* 796–802.

Gray, G. E., A. Paganini-Hill, R. K. Ross, and B. E. Henderson (1984). Assessment of three brief methods of estimation of vitamin A and C intakes for a prospective study of cancer: Comparison with dietary history. *Am. J. Epidemiol. 119,* 581–590.

Guthrie, H. A. (1984). Selection and quantification of typical food portions by young adults. *J. Am. Diet. Assoc. 84,* 1440–1444.

Hankin, J. H., G. G. Rhoades, and G. A. Glober (1975). A dietary method for an epidemiologic study of gastrointestinal cancer. *Am. J. Clin. Nutr. 28,* 1055–1060.

Hankin, J. H., A.M.Y. Nomura, J. Lee, T. Hirohata, and L. N. Kolonel (1983). Reproducibility of a dietary history questionnaire in a case-control study of breast cancer. *Am. J. Clin. Nutr. 37,* 981–985.

Heady, J. A. (1961). Diets of bank clerks: Development of a method of classifying the diets of individuals for use in epidemiologic studies. *J. R. Stat. Soc. (A) 124,* 336–361.

Hernandez-Avila M., C. Master, D. J. Hunter, J. Buring, J. Phillips, W. C. Willett, and C. H. Hennekens (1988). Influence of additional portion size data on the validity of a semi-quantitative food frequency questionnaire. *Am. J. Epidemiol. 128,* 891.

Howe, G. R., L. Harrison, and M. Jain (1986). A short diet history for assessing dietary exposure to N-nitrosamines in epidemiologic studies. *Am. J. Epidemiol. 124,* 595–602.

Humble, C. G., J. M. Samet, and B. E. Skipper (1987). Use of quantified and frequency indices of vitamin A intake in a case-control study of lung cancer. *Int. J. Epidemiol. 16,* 341–346.

Hunter, D. J., L. Sampson, M. J. Stampfer, et al. (1988). Variability in portion sizes of commonly consumed foods among a population of free-living women. *Am. J. Epidemiol. 127,* 1240–1249.

Marr, J. W. (1971). Individual dietary surveys: Purposes and methods. *World Rev. Nutr. Diet 13,* 105–164.

Morgan, R. W., M. Jain, A. B. Miller, et al. (1978). A comparison of dietary methods in epidemiologic studies. *Am. J. Epidemiol. 107,* 488–498.

Nichols, A. B., C. Ravenscroft, D. E. Lamphiear, and L. D. Ostrander Jr. (1976). Independence of serum lipids and dietary habits. *J.A.M.A. 236,* 1948–1953.

Pickle, L. W. and A. M. Hartman (1985). Indicator foods in vitamin A assessment. *Nutr. Cancer 7,* 3–23.

Rossi, P. H., and J. D. Wright (eds) (1983). *Handbook of Survey Research.* New York: Academic Press.

Samet, J. M., C. G. Humble, and B. E. Skipper (1984). Alternatives in the collection and analysis of food frequency interview data. *Am. J. Epidemiol. 120,* 572–581.

Sampson, L. (1985). Food frequency questionnaires as a research method. *Clin. Nutr. 4,* 171–178.

Stefanik, P. A. and M. F. Trulson (1962). Determining the frequency of foods in large group studies. *Am. J. Clin. Nutr. 11,* 335–343.

Stryker, W. S. (1987). Nutritional determinants of melanoma (thesis). Harvard School of Public Health.

Toth, B., D. Nagel, and A. Ross (1982). Gastric tumorigenesis by a single dose of 4-(hydroxymethyl) benzenediazonium ion of *Agaricus bisporis. Br. J. Cancer 46,* 417–422.

Wiehl, D. G. and R. Reed (1960). Development of new or improved dietary methods for epidemiological investigations. *Am. J. Pub. Health 50,* 824–828.

Willett, W. C., L. Sampson, C. Bain, et al. (1981). Vitamin supplement use among registered nurses. *Am. J. Clin. Nutr. 34,* 1121–1125.

Willett, W. C., L. S. Sampson, M. J. Stampfer, et al. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol. 122,* 51–65.

Willett, W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, and F. E. Speizer (1987). Dietary fat and the risk of breast cancer. *N. Engl. J. Med. 316,* 22–28.

Willett, W., L. Sampson, M. L. Browne, M. J. Stampfer, B. Rosner, C. H. Hennekens, and F. E. Speizer (1988). The use of a self-administered questionnaire to assess diet four years in the past. *Am. J. Epidemiol. 127,* 188–199.

Worsley, T. (1981). Psychometric aspects of language dependent techniques in dietary assessment. *Transactions of the Menzies Foundation 3,* 161–192.

6

Reproducibility and Validity of Food-Frequency Questionnaires

For reasons discussed in Chapter 5, the food-frequency questionnaire is usually the most appropriate method for dietary assessment in epidemiologic studies. It is, therefore, crucial to consider in detail the degree to which such questionnaires can measure true dietary intake. This chapter deals with approaches used to evaluate dietary questionnaires, the design of validation studies, and the analysis and presentation of data from validation studies.

In this chapter reproducibility refers to consistency of questionnaire measurements on more than one administration to the same persons at different times, realizing that conditions are never identical on repeated administration. Repeatability and reliability are frequently considered to be synonymous with reproducibility; however, the latter term has taken different meanings in other disciplines so that it is not used here. Validity refers to the degree to which the questionnaire actually measures the aspect of diet that it was designed to measure. This implies that a comparison is made with a superior, although always imperfect, standard. Reproducibility and validity can be addressed from several aspects. Most common is the relative ordering of subjects by the repeated measurements or different methods, which is typically evaluated by correlation coefficients. Comparisons of absolute levels can also be made, which usually involve examination of means and standard deviations; other methods of comparing measurements are discussed later in the chapter. The quantitative implications of different levels of reproducibility and validity are discussed in Chapter 12.

Because even subtle changes in the design of food-frequency questionnaires may affect their performance (see Chapter 5), each instrument should ideally be evaluated separately. Moreover, these structured questionnaires are culture-specific; even within a population they may perform differently among various demographic groups and subcultures. Thus, it is important to document the reproducibility and validity of any new questionnaire and to measure the performance of previously tested questionnaires for use in substantially different populations. As discussed in Chapter 2, an examination of dietary associations

with disease at the level of individual foods, food groups, and nutrients is frequently useful. Thus, studies of questionnaire reproducibility and validity at these three levels are important.

APPROACHES FOR EVALUATING DIETARY QUESTIONNAIRES

Various approaches that have been used to assess the performance of food-frequency questionnaires include the following:

1. Comparison of means
2. Proportion of total intake accounted for by foods included on the questionnaire
3. Reproducibility
4. Validity (comparison with an independent standard)
5. Comparison with biochemical markers
6. Correlation with a physiologic response
7. The ability to predict disease

Comparison of Means With Data From Other Sources

The comparison of mean nutrient intakes computed from a questionnaire with values derived from another source provides a simple and inexpensive method to assess performance. The comparison data may be external, such as the NHANES study (United States National Center for Health Statistics, 1979), or internal, based on another dietary assessment method among the same individuals who completed the questionnaire. For example, we have compared mean nutrient intakes based on a semiquantitative food-frequency questionnaire with means calculated from a 1-year diet record completed by 27 men and women (Willett et al., 1987; Table 6–1).

Although simple and inexpensive, a comparison of means provides limited information on validity. Similar mean values provide some reassurance that the questionnaire is reasonably comprehensive. It remains possible, however, that important items were not included on the questionnaire, but that the portion sizes assumed in the calculation of nutrient intakes were erroneously high; such compensating errors could produce correct mean values. Most seriously, such comparisons of means provides no information on the ability of the questionnaire to discriminate among persons.

Proportion of Total Intake Accounted for by Food Items Included on the Questionnaire

In developing a food-frequency questionnaire, one approach to the selection of foods for inclusion has been to use an open-ended method, such as 24-hour recall or diet record, to identify those foods that contribute importantly to absolute nutrient intake for a group as a whole (see Chapter 5). Such data have also been used as support for the completeness of the questionnaire.

Table 6–1. Comparison of mean nutrient intakes measured among 27 men and women using a 116-item food-frequency questionnaire, and 1-year diet record[a]

Nutrient	Diet record mean	Questionnaire mean
Total energy (kcal)	2,229 ± 706.9[b]	2,114 ± 1,012
Protein (g)	82.0 ± 24.8	87.0 ± 40.0
Total fat (g)	89.9 ± 30.1	81.9 ± 45.8
Saturated fat (g)	33.5 ± 13.1	31.9 ± 18.0
Linoleic acid (g)	14.0 ± 4.1	13.9 ± 7.8
Total carbohydrate (g)	258 ± 96	263 ± 116
Crude fiber (g)	4.4 ± 1.6	5.1 ± 2.8
Cholesterol (mg)	362 ± 122	332 ± 151
Oleic acid (g)	30.9 ± 9.85	29.4 ± 17.0
Vitamin A (IU)	6,434 ± 2,679	10,553 ± 6,194[c]
Niacin (mg)	21.9 ± 6.41	27.0 ± 12.2
Vitamin C (mg)	125 ± 87	146 ± 88
Calcium (mg)	894 ± 446	917 ± 586
Phosphorus (mg)	1,384 ± 504	1,420 ± 717
Thiamin (mg)	1.50 ± 0.55	1.30 ± 0.63
Riboflavin (mg)	1.91 ± 0.86	2.19 ± 1.31
Potassium (mg)	2,778 ± 1,045	3,076 ± 1,559
Iron (mg)	14.6 ± 5.92	13.6 ± 5.89

[a]Data were provided by 27 men and women aged 20 to 54.

[b]Mean ± standard deviation.

[c]Discrepancy for vitamin A is in part due to use of new USDA tables, which have dramatically changed vitamin A values for several vegetables. Use of older USDA values reduced this value to 8,511 IU.

From Willett et al., 1987.

As an example, we calculated the percentage of total nutrient intakes that were accounted for by food items on a compressed questionnaire (displayed in Appendix 5–1) and on a more comprehensive questionnaire that has been used in a wide variety of studies including, in slightly revised format, the 1984 Nurses Health Study (Willett et al., 1987a; Table 6–2). In this analysis, diet records completed by 194 women were used to compute the true total intake of 26 nutrients. Then we used the diet records to calculate the contribution of foods that were listed on the questionnaires to this true total nutrient intake; if all the foods recorded on the diet records had been listed on the questionnaire, the contribution would be 100 percent. In performing these calculations it quickly becomes apparent that there is not a one-to-one correspondence between foods on the questionnaire and foods reported in the diet record. For example, we had to collapse several hundred different codes for meat items on the diet records to correspond to the two meat items on the questionnaire. In addition, the coding of diet records and 24-hour recalls frequently requires the "dissection" of recipes into the basic ingredients such as the flour, shortening, and eggs used in baking. These basic ingredients would not be reported as such on the questionnaire, but would probably be recognized as the final products, such as cake or muffin. Because calculations like those shown in Table 6–2 include basic ingredients in the denominator (i.e., the total nutrient intakes from the diet records) but not in the numerator (i.e., the foods recorded in the diet records that are also listed

Table 6–2. Percentage of total nutrient intakes accounted for by foods listed on a 61-item compressed food-frequency questionnaire (see Appendix 5–1) and on an expanded 116-item revision of this questionnaire.

	Percentage of intake accounted for by foods on questionnaire	
Nutrient	Compressed questionnaire	Revised questionnaire
Total calories	69	93
Protein	77	95
Total fat	70	96
Saturated fat	75	96
Polyunsaturated fat	51	95
Monounsaturated fat	72	96
Cholesterol	85	97
Total carbohydrate	61	90
Crude Fiber	64	86
Sucrose	78	92
Total vitamin A	77	96
without supplements	73	95
Vitamin C	84	93
without supplements	76	90
Vitamin B_1	81	95
without supplements	67	91
Vitamin B_2	85	95
without supplements	75	92
Vitamin B_6	97	99
without supplements	84	95
Calcium	77	94
without supplements	77	94
Phosphorus (no supplements)	77	94
Potassium (no supplements)	73	93
Iron	75	93
without supplements	69	91
Mean	75	94

Percentages are underestimated because ingredients of recipes (e.g., flour, shortening) were included in the denominators but could not be attributed to specific foods. Nutrient intakes are based on four 1-week diet records completed by 194 women in 1980.

Adapted from Willett et al., 1987.

on the questionnaire), they will tend to underrepresent the proportions of nutrients accounted for by items on the questionnaire. Thus the data in Table 6–2 represent a conservative evaluation of the questionnaire by this particular criterion.

As a measure of performance, the approach of calculating the percentages of nutrients accounted for by a questionnaire is limited. A low percentage of nutrient intake accounted for would raise concern regarding the comprehensiveness of the form. As discussed in Chapter 5, however, such a questionnaire might still be reasonably discriminating if the foods had been carefully selected so as to explain maximally the between-person variation in nutrient intake. More seriously, a high percentage of nutrients accounted for does not guarantee validity, as the questionnaire may be inadequately interpreted by potential respondents. For example, it is tempting to use broad categories of foods on a questionnaire

such as "bread, crackers, and other baked goods." Although such questions account for a large percentage of absolute nutrient intake, they are likely to be more difficult to answer than a series of shorter specific questions. A series of more specific questions is thus likely to provide more accurate information, even though they may collectively account for a lower proportion of absolute nutrient intake.

The highest priority in designing a dietary questionnaire is usually to discriminate among persons with respect to their intake, rather than to estimate their absolute intake. Thus, it may be useful to employ diet record or short-term recall data to examine the proportion of between-person variance in specific nutrient intake that is accounted for by food items on a questionnaire. Although usually of more conceptual interest than an analysis based on absolute nutrient intakes, this computation requires a data set with many days of dietary intake for each subject. Stryker and colleagues (1987) have conducted such analyses for purposes of identifying items to be included on questionnaires. This approach, using open-ended dietary data, has apparently not been used to evaluate existing questionnaires.

Reproducibility

The reproducibility of questionnaire measurements made at two points in time can provide a useful first approximation of questionnaire performance. In conducting a reproducibility study, it is unrealistic to administer the questionnaire at a very short interval, such as a few days or weeks, as subjects may simply tend to remember their previous responses. When a longer interval of time is used, true changes in dietary intake, as well as variation in response, contribute to reduced reproducibility. Although it may be viewed as a disadvantage that such a measure of reproducibility reflects both the performance of a questionnaire and the true change in diet, both sources of variation realistically contribute to misclassification of long-term dietary intakes. Therefore, the difficulty in separating variation due to questionnaire performance from true change in diet is not extremely serious from the standpoint of evaluating measurement error. This feature, however, does hinder our capacity to measure the constancy of diet within individuals; the reproducibility correlation provides only an estimate of a lower limit of consistency. Assessment of the reproducibility of a method over several intervals of time may be useful in this respect. If a questionnaire refers to intake over the past year, administrations a few months apart should largely reflect variation associated with completing the questionnaire, whereas further decreases in reproducibility assessed at longer intervals, such as several years, should largely be due to true change in diet.

The reproducibility of food-frequency questionnaires has been examined under a wide variety of conditions (Table 6–3). In these studies, correlations have generally ranged from 0.5 to 0.7 for nutrient intakes measured at periods of 1 to 10 years (Rohan and Potter, 1984; Willett et al., 1985, 1987; Byers et al., 1987; Pietinen et al., 1988a, b). A notable exception is the study of Hankin and colleagues (1983), in which correlations for specific nutrients over an interval of only 3 months ranged from 0.12 to 0.41 among healthy whites. As the authors

suggested, this low level of reproducibility may be due to the questionnaire format in which the reference period of time was only 1 week. When assessed at an interval of 17 to 25 years, Byers and co-workers (1983) found the Spearman correlation for the reproducibility of a vitamin A index to be 0.29.

In a smaller number of studies, the reproducibility of specific food items has been examined (Table 6–4). Correlation coefficients have been considerably more variable than for nutrients. Among 323 U.S. men and women interviewed at an interval of 6 to 10 years, Byers and colleagues (1987) found average correlations of 0.41 for vegetables, 0.41 for fruits, 0.53 for diary products, and 0.39 for meats. Colditz and co-workers (1987) compared frequencies of foods reported by 1497 women at an interval of approximately 9 months. Correlations were highest for beverages ($r = 0.70$) and ranged from 0.60 to 0.70 for foods eaten frequently and from 0.34 to 0.45 for foods eaten infrequently. In this study, the reproducibility of food intake did not vary appreciably by age, relative weight, cigarette smoking status, or alcohol intake.

Correlation coefficients on the order of 0.5 to 0.7, which appear to be typical for the reproducibility of nutrient intakes, may seem disappointingly low for those accustomed to the reproducibility of laboratory measurement made under highly controlled conditions. Nevertheless, this level of reproducibility is comparable with that of many biological measurements made among free-living subjects over a period of months or years. For example, measurements of serum cholesterol and blood pressure have similar degrees of reproducibility (Table 6–5) and yet are strong and consistent predictors of disease in epidemiologic studies.

The interpretation of reproducibility studies should be somewhat asymmetric: a low degree of reproducibility is a definite indication that the questionnaire does not provide a valid measure of long-term intake. On the other hand, a high degree of reproducibility does not ensure validity as high correlation can be simply the result of correlated error (i.e., systematic within-person error). For example, a questionnaire that has omitted important sources of a nutrient or that includes questions that are consistently misinterpreted may be highly reproducible, but fail to provide a true measure of intake for that nutrient. Because reproducibility studies are usually quick and inexpensive to conduct, they are an appropriate part of the questionnaire evaluation, but cannot substitute for studies of validity.

Validity: Comparison of Individual Values With an Independent Measure of Diet

The ability of a questionnaire to discriminate among individuals is most directly evaluated by comparing individual estimates of nutrient intake based on the questionnaire with those measured by a more accurate method, that is, a gold standard. It has been frequently said that there is no perfect measure of dietary intake, with the implication that validation studies are not possible. Lack of a perfect standard is, however, not unique to dietary intake; all measurements have error, although these errors differ in their magnitude. Thus, validation studies never compare an operational method with absolute truth, rather they

Table 6–3. Reproducibility of nutrient intakes measured by repeated food-frequency questionnaires

Source	Population	FFQ design	Time to complete FFQ (min)	Interval	Range of correlations
Hankin et al. (1983)	Japanese–Hawaiian women; cases, controls (n = 117)	43 items, portions estimated by pictures, interview, focus on fat, cholesterol, and protein	NI	3 mo	0.12 protein to 0.41 total fat
Byers et al. (1983)	Men and women patients admitted to Roswell Park Memorial, 50–74 yr (n = 175)	12 items, 12 response categories, interview focus on vitamin A intake	15	17–25 yr	−0.14 bread to 0.52 coffee, 0.29 vitamin A
Rohan and Potter (1984)	South Australian men and women, 35–78 yr (n = 70)	141 items, standard portions, first questionnaire by interview, second by mail	NI	3 yr	0.25 protein to 0.87 alcohol (men) 0.43 calcium to 0.79 alcohol (women)
Willett et al. (1985)	Registered nurses from Mass., 34–59 yr (n = 194)	61 items, 9 response categories, standard portions, mailed, focus on nutrients related to cancer	15	9–12 mo	0.52 vitamin A without supplements to 0.71 sucrose

Willett et al. (1987)	Registered nurses from Mass., 39–63 yr (n = 150)	116 items, 9 response categories, standard portions, mailed, focus on nutrients related to cancer	25	3 yr	0.44 total carbohydrate to 0.62 vitamin C
Byers et al. (1987)	U.S. men and women (n = 323)	129 foods plus servings sizes, interview (reinterviewed for 47 foods only)	>60 min	6–10 yrs	0.50 fat to 0.61 vitamin A and fiber
Colditz et al. (1987)	U.S. nurses (n = 1497)	61 item form vs. 116 item form	15 and 25 min	9 mo	0.40 *trans*-fatty acids to 0.71 vitamin E
Pietinen et al. (1988a)	Finnish men (n = 121)	44 foods, frequency only	NI	6 mo	0.53 vitamin A to 0.85 polyunsaturated fat
Pietinen et al. (1988b)	Finnish men (n = 121)	276 foods and mixed dishes, pictures for portion review with dietician	2 hr +½ hr for review	6 mo	0.54 vitamin A to 0.74 sucrose and 0.88 alcohol

FFQ = food-frequency questionnaire, NI = no information

Table 6–4. Reproducibility of food intake measured by repeated food-frequency questionnaires

Source	Population	FFQ design	Time required to complete FFQ (min)	Interval between FFQs	Rank order correlations or % agreement
Acheson and Doll (1964)	GI cancer cases, control men and women over 75 yr (n = 63)	56 items, 5 response categories, interviews	NI	3 mo	90% agreement within 1 category
Graham et al. (1967)	Gastric cancer cases, control subjects (n = 99)	27 items, 4 response categories, interview, focus on Polish–American diets	NI	18 mo	81% agreement for exact category
Nomura et al. (1976)	Japanese–Hawaiian men (n = 109)	33 items, portions estimated by pictures, focus on food associated with GI cancer	15	6 mo–2 yr	−0.05 dried fish, 0.03 sausage to 0.56 tomato juice, 0.71 coffee
Byers et al. (1987)	U.S. men and women (n = 323)	129 foods plus serving sizes, interview (reinterview for 47 foods only)	>60 min	6–10 yr	0.18 for roast beef to 0.71 for coffee
Colditz et al. (1987)	U.S. nurses (n = 1497)	61 item form vs. 116 item form	15 and 25 min	9 mo	0.70 for beverages, 0.60–0.70 for frequently eaten foods 0.34–0.45 for infrequently eaten foods
Thompson et al. (1987)	U.S. men and women (n = 1184)	83 items, 8 responses categories with "card sort" response	NI	15 yr	average intraclass r for 77 nonseasonal foods was 0.37 for men and 0.34 for women

FFQ = food-frequency questionnaire, GI = gastrointestinal, NI = no information.

Table 6–5. Reproducibility of commonly used epidemiologic variables

Study	Variables	Population	Interval	Correlation
Shekelle et al. (1981)	Serum cholesterol	1900 men	1 yr	0.65
Rosner et al. (1977)	Systolic blood pressure	863 men and women	4 yr	0.64
	diastolic blood pressure			0.60
Willett et al. (1983)	Plasma retinol	15 men and women	8 wk	0.58
	plasma total carotenoid		8 wk	0.60
	plasma alpha-tocopherol		8 wk	0.50
Gordon and Shurtleff (1973)	Blood glucose	1597 men	2 yr	0.58
		1841 women		0.52
	Vital capacity	2085 men	2 yr	0.79
		2603 women		0.77
	Pulse rate	2008 men	2 yr	0.52
		2529 women		0.49

compare one method with another method that is judged to be superior. Given that neither method is perfect, it is crucial that the errors of both methods be as independent (i.e., uncorrelated) as possible to avoid spuriously high estimates of validity. For example, in some studies a dietary questionnaire has been compared with a detailed diet history interview. This provides a very limited assessment of validity as the major sources of error in the questionnaires (e.g., memory, interpretation of questions, and perception of serving sizes) are likely to be replicated in this diet history. To the extent that errors in the comparison method are uncorrelated with error in the questionnaire being evaluated, the correlation between the two tends to be underestimated. The lack of a perfect comparison method also indicates a continued need to search for better gold standards.

Among the available and feasible comparison methods for validating a food-frequency questionnaire, diet records are likely to have the least correlated errors. Major sources of error associated with food-frequency questionnaires are due to the restrictions imposed by a fixed list of foods, memory, perception of portion sizes, and interpretation of questions. These sources of error are minimally shared by diet records as they are open-ended, they do not depend on memory (foods are recorded on a meal-by-meal basis), they allow direct assessment of portion sizes by measurement of weight or dimensions, and errors related to interpretation are largely at the level of the dietitian coding the records rather than the subject. Because errors associated with food-frequency questionnaires and diet records are largely independent, validity, if anything, tends to be understated. When used as a standard to assess questionnaire validity, diet records should, in principle, be kept for a sufficient number of days to represent average intake and cover the interval of time corresponding to the questionnaire, typically 1 year.

One source of error that is likely to remain correlated when comparing nutrient intakes measured by a food-frequency questionnaire with diet records is the food composition data. Nutrient intakes calculated from the two methods are usually based on a similar body of published data. Thus, for nutrients that vary

greatly for different examples of the same food, the calculated values from the diet record may be incorrect, but still correlated with the questionnaire. For example, intakes of selenium (which can vary widely depending on the soil content where the food was produced) or folic acid intake (which is sensitive to processing, cooking, and storage) may appear to be measured well as judged by comparison of a questionnaire with a diet record, but both methods may be incorrect due to the nutrient composition tables being unrepresentative of the foods that were actually consumed. As noted in Chapter 4, the process of keeping a diet record may alter food intake; to the extent that this is a departure from usual food habits (which are the focus of a food-frequency questionnaire), this will also tend to reduce the correlation between the questionnaire and the record.

The primary alternative to the use of diet records as a standard for evaluating a food-frequency questionnaire is the collection of multiple 24-hour recalls. Because errors are more likely to be correlated with use of this method (both rely on memory and perception of serving sizes), it is probably suboptimal. In many situations, however, such as when subjects are illiterate or less than highly motivated, multiple 24-hour recalls may be the only reasonable option.

The validity of food-frequency questionnaires has been examined in a number of studies (Table 6–6). In several early investigations, usually based on a relatively small number of subjects, impressively high correlations for both food and nutrients were observed between simple dietary questionnaires and diet histories (Abramson et al., 1963; Browe et al., 1966; Balogh et al., 1968; Epstein et al., 1970). Stefanik and Trulson (1962) found that a simple food-frequency questionnaire could reasonably characterize cultural differences in eating habits between Irish-born and Italian-born American men.

More recently, Jain and colleagues (1982) compared a simple, mailed questionnaire with a detailed diet history interview conducted after an interval of 3 months among 50 women. Reasonably high correlations were observed, ranging from 0.47 for cholesterol to 0.72 for vegetable protein. Lower correlations were observed by Gray and co-workers (1984) for intake of vitamins C ($r = 0.29$) and A ($r = 0.03$) when the diet history was completed at an interval of 15 months. The authors attributed these low correlations to the broad response categories that were used. In this study the majority of important food sources of these vitamins were used one to several times a week; however, the questionnaire required subjects to describe these frequencies as either "a few times a week" or "daily or almost daily." The apparently carefully conducted study by Stuff and co-workers (1983) is a clear contrast to the generally encouraging findings of others. Among 40 lactating women, these investigators found essentially no correlation between nutrients measured by a food-frequency questionnaire and a 1-week diet record completed the following week. There is no obvious explanation for this discrepant finding, although it might possibly relate to unusual variation in diet during lactation.

Among the most detailed studies was our validation assessment of a 61-item food-frequency questionnaire used in 1980 as part of the Nurses Health Study (Willett et al., 1985). In this study, women, aged 34 to 59 years, living in the greater Boston area, were randomly selected from among the respondents to the

1980 Nurses Health Study questionnaire. Approximately 3 months after completing the questionnaire, each participant was provided with a dietetic scale and instructed by a research dietitian in recording food intake for 1 week. The process of keeping a 1-week diet record was repeated at 3-month intervals for a total of 4 weeks. At the end of either the third or fourth week, a second food-frequency questionnaire, identical to the first, was completed (Fig. 6–1).

Because we inquired on the dietary questionnaire about the average intake of foods over the past year, the Nurses Health Study validation study was designed to measure optimally food over a 1-year period using diet records. Thus, the repeated diet record measurements at 3-month intervals should have accounted for seasonal changes and medium-term drifts in food habits over this period of time. Because day-to-day variability in intake of many nutrients is large, we collected 28 days of intake data per person to provide a reasonably stable estimate of true long-term intake for nearly all nutrients. Beaton and colleagues (1983), however, have suggested that even this number of days of diet recording may be insufficient to measure accurately intake of some nutrients such as vitamin A; our analyses have indeed indicated that the correlations with highly variable nutrients (e.g., cholesterol intake) based on 28 days of data are still somewhat attenuated due to within-person variability (see Chapter 12).

In designing the Nurses Health Study food-frequency questionnaire validation study, we chose to administer the questionnaire before and at the end of the diet record collection out of concern that the process of recording diet might alter awareness of food intake and thus artificially improve accuracy in completing the questionnaire. Although this possibility is avoided by comparing the initial questionnaire with the diet records collected during the subsequent year, this comparison would tend to underestimate validity as the initial questionnaire asked about diet during the year before recording. The use of a questionnaire both before and after recording provides, in some sense, minimal and maximal estimates of true validity. An added benefit of administering the questionnaire before and after the diet record collection is the opportunity to access the reproducibility of the questionnaire over that 1-year period. Reassuringly, correlations betwen the initial food-frequency questionnaire and the mean diet record intake were only slightly lower than those between the repeat questionnaire and the diet record mean, which represents the conceptually appropriate

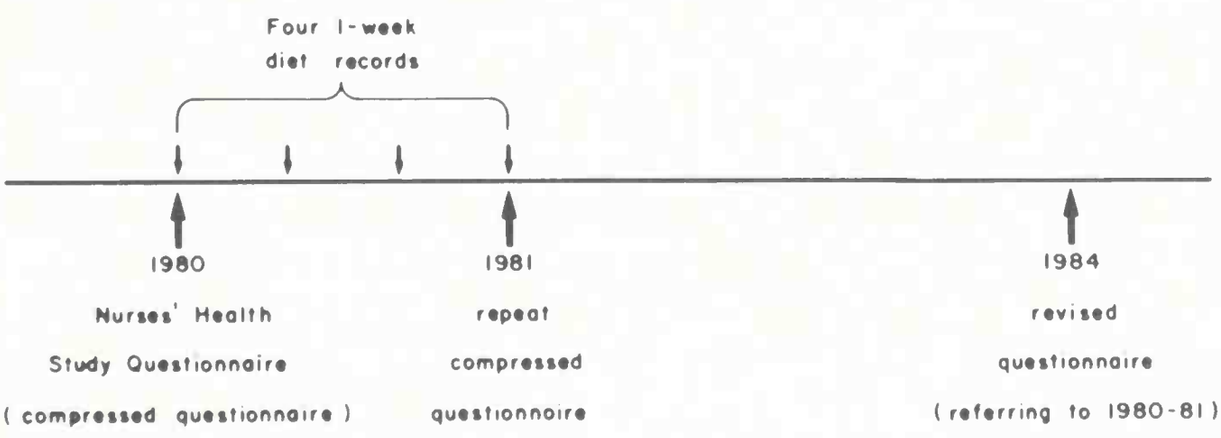


Figure 6–1. Design of Nurses' Health Study food-frequency questionnaire validation study. (From Willett et al, 1985; reproduced with permission.)

Table 6–6. Comparison of food-frequency questionnaires with other dietary assessment methods[a]

Source	Population	Comparison methods	Interval between methods	Reference period	Range of correlations	Comments
Browe et al. (1966)	Albany Cardiovascular Health Study (n = 29 men)	Diet history	4 wk	Usual monthly meal patterns	0.66 vegetable protein to 0.79 cholesterol	
Abramson et al. (1963)	Pregnant Israeli women (n = 60)	30 min interview	0	Usual diet for previous 3 mo	0.42 bread and rolls to 0.99 herring	Correlations were lower for groups of foods
Balogh et al. (1968)	Israeli men (n = 48)	Diet history	A few hours	Usual diet	0.78 polyunsaturated fatty acids to 0.95 animal protein	
	(n = 14)	Food record		Week after FFQ	0.69 polyunsaturated fatty acids to 0.94 total fat	
Epstein et al. (1970)	Heterogenous Israeli adults (n = 161)	Diet history	4–30 days	Usual diet	0.14 linoleic acid to 0.76 total protein	
Hankin et al. (1975)	Japanese–Hawaiian men (n = 50)	Food record	0	1 wk FFQ recalled previous 1-week record	0.47 beef to 0.88 shrimp, raw fish, and coffee	
Jain et al. (1982)	Canadian women (n = 50)	Diet history	3 mos	Usual diet for previous 2 mo FFQ reports usual diet	0.47 cholesterol to 0.72 vegetable protein	

Stuff et al. (1983)	Lactating Texan women (n = 40)	Food record	0	FFQ reports usual diet	0.00 iron and phosphorous to 0.24 calcium	FFQ requested usual intake during a time that diet might be changing
Willett et al. (1985)	Registered nurses (n = 194)	Diet record	1 mo–1 yr	Previous year	0.36 Vitamin A without supplements to 0.75 vitamin C	
Gray et al. (1984)	Southern California elderly (n = 50)	Diet history	15 mo	Usual diet	0.16 vitamin A to 0.36 vitamin C	Very broad response categories were used in FFQ
Hunt et al. (1979)	California adults (n = 46)	24 hr recall	1–6 wk	Usual diet	0.04 vitamin A to 0.61 carbohydrate	Unusually large portion sizes for vitamin C foods were used in FFQ
Jensen et al. (1984)	Danish men and women (n = 79)	Dietary household survey	15–25 yr	15–25 yr earlier	0.13 fresh vegetables to 0.42 energy	Correlations similar for present vs. remote diet
Willett et al. (1987)	U.S. Registered Nurses, 39–63 yr (n = 150)	Four 1-wk diet records	3–4 yr	1 yr (3–4 yr earlier)	0.28 iron to 0.61 carbohydrate	Correlations mainly 0.50 or greater
Pietinen et al. (1988a)	Finnish men (n = 189)	Twelve 2-day diet records (vs. 44-item questionnaire)	1–6 mo.	1 yr	0.38 vitamin A to 0.69 polyunsaturated fat	Adjustment for energy increased correlations for most nutrients
Pietinen et al. (1988b)	Finnish men (n = 189)	Twelve 2-day diet records (vs. 273-item questionnaire)	1–6 mo	1 yr	0.51 vitamin A to 0.73 polyunsaturated fat and 0.80 alcohol	Adjustment for energy had little effect on correlations

[a]Extension of data compiled by Sampson, 1985.

Table 6–7. Pearson correlation coefficients for comparison of semiquantitative food-frequency questionnaire scores with the means of four 1-week diet records, both unadjusted and adjusted for total caloric intake

	Questionnaire 1 vs. means of 4 records		Questionnaire 2 vs. means of 4 records	
Nutrient[a]	*Unadjusted*	*Adjusted*[b]	*Unadjusted*	*Adjusted*[b]
Protein	0.18	0.37	0.33	0.47
Total fat	0.27	0.48	0.39	0.53
Saturated fat	0.31	0.49	0.44	0.59
Polyunsaturated fat	0.31	0.42	0.40	0.48
Cholesterol	0.46	0.61	0.52	0.61
Total carbohydrate	0.48	0.44	0.53	0.45
Sucrose	0.52	0.41	0.60[c]	0.54[c]
Crude fiber	0.43	0.51	0.46	0.58
Total vitamin A	0.37	0.43	0.41	0.49
without supplements	0.21	0.28	0.26	0.36
Vitamin B_6	0.44	0.47	0.54[c]	0.58[c]
without supplements	0.32	0.43	0.43[c]	0.54[c]
Vitamin C	0.53	0.56	0.73[c]	0.75[c]
without supplements	0.46	0.52	0.63[c]	0.66

[a]Nutrient intakes transformed using $\log_e$ to improve normality.

[b]Intakes adjusted using the residuals from regression models with caloric intake as the independent variable and nutrient intakes as the dependent variables.

[c]$p < 0.05$ for comparison of Pearson r for questionnaire 2 vs. means of four records with Pearson r for questionnaire 1 vs. means of four records.

Data based on information provided by 173 female registered nurses aged 34–59 years and residing in the Boston area, 1980–1981.

From Willett et al., 1985.

time sequence (Table 6–7). In these data we found that adjustment for total caloric intake improved the correlations for macronutrients, presumably due to "canceling" of correlated errors (see Chapter 11 for a more detailed discussion). After adjustment for caloric intake, correlations ranged from 0.35 for vitamin A without supplements to 0.75 for vitamin C with supplements.

A recent large study conducted in Finland by Pietinen and co-workers (1988a, b) provided the opportunity to compare the relative validity of an extremely short questionnaire and an extremely detailed questionnaire, which were both self-administered. The short questionnaire consisted of 44 foods and only information on simple frequency-of-use was obtained. The detailed questionnaire contained 273 foods and mixed dishes, and a corresponding 63-page book of portion sizes; this required 2 hours to complete and ½ hour to review with a nutritionalist. When compared with the mean of twelve 2-day diet records, reasonable correlations were seen with both methods (Table 6–6); after adjustment for total energy intake, the correlations with the diet record were on average only 0.07 higher for the detailed questionnaire compared with the extremely short form. Because few study populations would tolerate a more extensive questionnaire than the detailed one used in this study, these data provide an indication of the limits of the food-frequency methodology and demonstrate the rapidly decreasing marginal gain in information obtained with increasingly detailed questionnaires.

Most validation studies have been analyzed on the basis of nutrient intakes; however, it is also possible to compare methods at the level of individual foods or food groups. Such comparisons may be particularly helpful in focusing attention on the questionnaire items that are performing poorly, thus suggesting specific areas that could be improved in subsequent questionnaires.

Data from the 1980 Nurses Health Study dietary questionnaire validation analysis were also used to examine the validity of specific food items (Salvini et al., 1989). In conducting a food-based validation analysis, it quickly becomes apparent that the task is complex and laborious, as there is not a direct one-to-one correspondence between items on the questionnaire and items in the food record. In general, foods are recorded in much more detail in the diet records and sometimes in the form of complex mixed dishes. For example, to compare beef, pork, and lamb intake assessed by the two methods, it was necessary to combine hundreds of items representing different types and forms of meat and dissect dozens of different mixed dishes. Furthermore, day-to-day variation in intake of most specific foods is substantially greater than variation in nutrients. Thus, even 28 days of diet recording did not adequately represent usual intake of many foods. For this reason, correlation coefficients were corrected for within-person variation in the diet record (see Chapter 12).

Data for selected foods analyzed in this manner using the repeat questionnaire (Fig. 6–1) are presented in Table 6–8. Correlations for most foods were similar in magnitude to those observed for nutrient intakes, indicating that the questionnaire performed reasonably well. The questionnaire, however, substantially overestimated intake of "spinach and other greens" and correlated poorly with the diet record for this item. We suspected that this problem arose because of deletion of lettuce from the questionnaire in the process of compressing the forms so that individuals variably considered lettuce as "other greens." This was addressed in subsequent questionnaires by deleting the phrase "other greens" (which leaves much to individual interpretation) and adding back items on let-

Table 6–8. Comparison of average intake of selected specific foods reported on a compressed 61-item questionnaire with intake measured by 28 days of dietary record among 173 U.S. women (unpublished data)

	Diet record mean	Questionnaire mean	Records vs. questionnaire (Pearson correlation)	
			Crude	*Adjusted*
Low fat milk (cups)	0.28	0.53	0.79	0.81
Whole milk (cups)	0.27	0.22	0.62	0.62
Margarine (pats)	1.24	1.50	0.71	0.76
Butter (pats)	0.97	0.64	0.79	0.85
Spinach, other greens (½ cup)	0.06	0.28	0.08	0.17
Broccoli (½ cup)	0.07	0.17	0.49	0.69
Apples (1 fruit)	0.20	0.33	0.66	0.80

Variables transformed by $\log_e$ to improve normality. Adjusted correlations are corrected for within-person variation in dietary record intake (see Chapter 12).

tuce to the questionnaire. When responses by the same women to the 1982 and 1984 Nurses Health Study questionnaire were compared with the diet records collected in 1980, the correlations for spinach were substantially higher and similar to those for other foods, indicating that the altered wording had corrected the problem in the original questionnaire.

As noted earlier, the interpretation of correlation coefficients on the order of 0.5 to 0.6 typically seen in validation studies may be somewhat difficult for scientists accustomed to correlations of 0.90 or larger when laboratory methods are compared. Although the impact of this degree of error on disease relationships is discussed quantitatively in Chapter 12, the validation study described previously provided the opportunity to compare the relative capacities of the simple questionnaire with a 1-week diet record to represent long-term intake (Table 6–9). The correlations of both a single 1-week (week 4) diet record and the second questionnaire with the mean of the first 3 weeks of diet recording were examined. Although the correlations were slightly higher between the fourth diet record and the first three diet records than between the questionnaire and the first three diet records, the differences were generally small, indicating that the simple questionnaire captured nearly as much information on individual dietary intakes as the considerably more expensive diet record.

The existence of the four 1-week diet records collected in 1980 and 1981 also provided the opportunity to examine the validity of the food-frequency ques-

Table 6–9. Comparison of the second semiquantitative food-frequency questionnaire and diet record four nutrient scores with mean nutrient scores from diet records 1–3

	Pearson correlation coefficients for calorie-adjusted intakes	
Nutrient[a]	*Questionnaire 2 vs. records 1–3*	*Record 4 vs. records 1–3*
Protein	0.48	0.66[b]
Total fat	0.52	0.64
Saturated fat	0.58	0.64
Polyunsaturated fat	0.46	0.60[b]
Cholesterol	0.58	0.56
Total carbohydrate	0.42	0.76[b]
Sucrose	0.56	0.63
Crude fiber	0.56	0.78[b]
Total vitamin A	0.49	0.63[b]
without supplements	0.36	0.49
Vitamin B_6	0.59	0.80[b]
without supplements	0.51	0.62[b]
Vitamin C	0.76	0.75
without supplements	0.64	0.72

[a]Nutrients transformed using $\log_e$ to improve normality.

[b]$p < 0.05$ for comparison of Pearson r for record 4 versus records 1–3 with Pearson r for questionnaire 2 versus records 1–3.

(Data based on information provided by 173 female registered nurses aged 34–59 years and residing in the Boston area, 1980–1981).

From Willett et al., 1985.

tionnaire for assessing diet several years in the past (Willett et al., 1988). In 1984 we mailed a revised and expanded (116-item) version of the semiquantitative food-frequency questionnaire to the women who had recorded their diet in 1980 to 1981; the questionnaire was worded to ask about food intake 3 to 4 years earlier. This expanded questionnaire included additional foods that had been eliminated during the development of the 1980 questionnaire due to their small independent contribution to between-person variance in nutrient intakes (see Chapter 5). The revised form thus accounted for a larger absolute intake of nutrients (see Table 6–2). In addition, a number of foods that had been previously asked as a single broad item were separated into discrete items. Mean nutrient intakes assessed by the revised questionnaire were similar to those measured by the diet record (see Table 6–1). Correlations with intake measured by diet records again tended to be higher after adjustment for total caloric intake in the case of macronutrients and were, overall, quite similar to correlations observed for the questionnaire completed at the end of the diet recording in 1980 to 1981 (Tables 6–8 and 6–10). Because memory inevitably tends to fade with time (Bradburn et al., 1987), this similarity in correlations was interpreted as meaning that the revised questionnaire represented an improvement that was sufficient to compensate for an increased interval of time. These data also provide evidence that the semiquantitative food-frequency questionnaire method can reasonably measure dietary intake several years earlier, as would typically be done in a case-control study. To some extent, the recall of past diet may be a function of some consistency in diet over time and the tendency of recent food intake to influence the recall of past diet (see Chapter 7); these data cannot distinguish this effect.

Although the design of the validation study we employed had many attractive features, it was an expensive process. The full-time efforts of a research dietitian were required for over 2 years in addition to the support of programmers and data-entry personnel. Although improved computer-based systems for the coding of diet records enhance efficiency to some extent, the cost of such a study may sometimes be difficult to justify when evaluating further versions of a dietary questionnaire or applications in different populations. In such instances, most of the useful information appears to be obtainable with a substantially smaller number of days of diet recording or 24-hour recalls (discussed later and in Chapter 12).

In principle, the use of direct observation of food intake would provide an excellent gold standard for validation of a dietary questionnaire. Unfortunately, this approach is usually limited, as such observations require an artificial environment. Direct observation was recently used to assess the validity of reported food intake during a specific potluck luncheon, which could be an appropriate reference period for a foodborne outbreak (Decker et al., 1986). During the luncheon, a video recorder and camera were used to record the activities of all 32 attendees under the guise that new equipment was being tested. Two to three days later attendees were asked to complete a questionnaire regarding the use or nonuse of all foods that were served, and these responses were compared with the videotaped actual food consumption. Although the associations between reported and recorded intakes were high, only 10 of the 32 subjects made no

Table 6–10. Correlations between nutrients estimated by a 116-item semiquantitative food-frequency questionnaire and the average of four 1-week diet records collected 3–4 years earlier

	1984 revised questionnaire vs. 1980 diet records	
Nutrient	*Crude*	*Adjusted*
Total calories	0.37	—
Protein	0.29	0.52
Total fat	0.37	0.54
Saturated fat	0.38	0.52
Polyunsaturated fat	0.43	0.58
Monounsaturated fat	0.37	0.48
Cholesterol	0.50	0.57
Total carbohydrate	0.47	0.61
Crude fiber	0.43	0.56
Sucrose	0.50	0.45
Total vitamin A	0.36	0.44
without supplements	0.26	0.37
Vitamin C	0.38	0.54
without supplements	0.50	0.54
Vitamin B_1	0.32	0.58
without supplements	0.46	0.55
Vitamin B_2	0.33	0.57
without supplements	0.37	0.52
Vitamin B_6	0.50	0.54
without supplements	0.39	0.57
Calcium	0.56	0.56
without supplements	0.46	0.51
Phosphorus (no supplements)	0.32	0.51
Potassium (no supplements)	0.41	0.53
Iron	0.55	0.55
without supplements	0.29	0.28

Diet record data were based on means of four 1-week records. All data were transformed by $\log_e$ to improve normality.

Data are based on responses of 150 women.

From Willett et al., 1988.

error. Mullen and colleagues (1984) used direct observation of food consumption on multiple days by students using a cafeteria to compare with their responses to a food frequency questionnaire. Unfortunately, this carefully collected data set was analyzed by comparing the frequencies of use (assessed by the two methods) for specific foods *within* each subject, thus providing no information on the capacity of the questionnaire to distinguish among subjects.

Comparison With a Biochemical Indicator of Dietary Intake

Biochemical indicators of dietary intake have great intuitive appeal as the gold standard to assess the validity of a dietary questionnaire. The fundamental advantage in using a biochemical indicator is that measurement errors should be essentially uncorrelated with errors in any dietary questionnaire. Thus, the

capacity to demonstrate a correlation between the questionnaire assessment of intake and the biochemical indicator provides almost unquestionable qualitative documentaion of validity. Although biochemical markers are undeniably attractive as standards for studies of validity, they have many limitations (see Chapter 9). Biochemical indicators are unlikely to be influenced by dietary intake alone as individuals generally differ to some degree in the absorption and postabsorptive metabolism of most nutrients. These differences cause variability in the biochemical indicator unrelated to intake. Furthermore, other sources of physiologic variation, such as levels of binding proteins or poorly understood fluctuations related to diurnal or menstrual cycles, may influence the biochemical level.

Variation in dietary intake from day-to-day will itself cause fluctuation in biochemical indicator levels; the effect of this source of variation differs among indicators depending on their time-integrating capacity. To the extent that the biochemical indicator is serving as a standard for long-term dietary intake, such variation is a source of error. Finally, the technical error associated with laboratory measurement contributes to variation in the biochemical indicator. The net effect of all these sources of variation is to weaken correlations between a dietary questionnaire used to assess long-term intake and the biochemical indicator. For this reason it is likely that the magnitude of such correlations will tend to be rather modest, even when the dietary measurements are highly accurate and precise. Differences in the bioavailability of nutrients from food to food could cause further weakening of these correlations; however, this source of error is probably justifiably attributed to the questionnaire as it is related to inappropriate food composition data.

Several strategies can be employed to account for extraneous variation in the biochemical standard used in a validation study. To the extent that factors causing this variation can be measured, adjustment can be made for them, thus improving observed correlations. For example, vitamin E is nonspecifically carried in lipoprotein particles and is, therefore, strongly correlated with serum cholesterol; adjusting serum vitamin E levels for serum cholesterol thus improves the correlation with both intake (Willett et al., 1983; Stryker et al., 1988) as well as with the availability to functionally relevant tissues (Sokol et al., 1984). Attenuation of correlation coefficients due to true day-to-day variation or random laboratory error can be overcome by either obtaining a large number of replicate samples per subject or by using a smaller number of replicates and correcting statistically for random error (see Chapter 12).

In addition to the limitation that biochemical indicators do not provide a pure reflection of diet, their use is markedly constrained by the lack of any marker for many dietary factors of major interest. For example, biochemical indicators do not exist for intake of total fat, total carbohydrate, sucrose, or fiber. For some other nutrients, biochemical measurements do exist, but homeostatic regulation is so strong (and thus their association with intake so weak) that they are of minimal use as standards in a validation study within the range of typical diets. Such highly regulated nutrients include plasma levels of cholesterol, retinol, and calcium. In spite of these limitations, the use of biochemical markers in a validation study can be helpful when they are available. For example, a null

Table 6–11. Comparison of food-frequency questionnaires with biochemical indicators of diet

Nutrient intake	Subjects	Biochemical measure	Correlation(r)
Dietary carotene (Willett et al., 1983)	59 men and women	Plasma total carotenoid	0.35[a]
Dietary vitamin E (Willett et al., 1983)	59 men and women	Plasma alpha-tocopherol	0.34[a]
Dietary carotene (Stryker et al., 1988)	137 men/193 women	Plasma beta-carotene	0.36/0.42[a]
Dietary vitamin E (including supplements) (Stryker et al., 1988)	137 men/193 women	Plasma alpha-tocopherol	0.53/0.51[a]
Dietary carotene (Russell-Briefel et al., 1985)	187 men	Plasma total carotenoid	0.25
Vitamin B_6 (Willett, 1985)	94 men/25 women	Plasma pyridoxal phosphate	0.37/0.39
Saturated fat[b] (Sacks et al., 1986)	19 men and women	Serum cholesterol	0.54
Linoleic acid[b] (Sacks et al., 1986)	19 men and women	Serum choleserol	−0.51
Omega-3 fatty acids (Silverman et al., 1988)	42 men	Plasma phospholipid eicosapentaenoic acid	0.54

[a]Adjusted for caloric intake and plasma lipids.

[b]Subjects were participants in a dietary intervention trial: change in diet was correlated with change in serum cholesterol.

association between a nutrient intake measured by questionnaire and the risk of a disease is more meaningful if the questionnaire measurement has been shown to be correlated with a biochemical indicator.

Relatively few examples exist of questionnaire validations using biochemical measurements (Table 6–11). As part of the baseline data in a supplementation trial (Willett et al., 1983), we measured intake of dietary carotene (more specifically, carotenoids with provitamin A activity) and vitamin E, as well as plasma levels of total carotenoids and alpha-tocopherol (the major form of vitamin E). After adjusting nutrient intakes for total caloric intake and plasma nutrients for plasma lipid values, the correlations were 0.35 for carotene and 0.34 for vitamin E.

In a larger population, using a slightly more complete questionnaire to assess dietary carotene and a measure of plasma beta-carotene rather than total carotenoid, Stryker and co-workers (1988) observed a correlation of 0.36 for men and 0.41 for women between intake and plasma level. Cigarette smoking appeared to alter the relationship between carotene intake and blood level; among nonsmokers the correlation was 0.44 for men and 0.45 for women. In the same study the correlation for vitamin E including supplements was 0.55 for men and 0.51 for women; when supplements were not included the correlation was lower. Vitamin B_6 assessed by a semiquantitative food-frequency questionnaire and plasma pyridoxal phosphate (a measure of vitamin B_6 status) were measured

among community controls in a study of myocardial infarction; correlations between the two were 0.37 for men and 0.39 for women (Willett, 1985).

Known biochemical responses to a dietary factor can also serve as useful markers in validation studies. For example, in many studies, it has been shown that increases in alcohol consumption produce an elevation in plasma high-density lipoprotein cholesterol (HDL-C) levels. Thus, the demonstration of a correlation between a questionnaire measure of alcohol intake and HDL-C provides qualitative evidence that alcohol intake is being measured with at least some degree of validity. Because other factors affect HDL-C levels, this cannot provide a highly quantitative measure of misclassification. This approach, however, can also be used to estimate the validity of one method relative to an other. For example, in the Nurses Health Study validation substudy (Willett, 1987b), we found that the correlation between alcohol intake measured by food-frequency questionnaire and HDL-C was as strong or stronger than the correlation between alcohol intake measured by 28 days of diet record and HDL-C (Figure 6–2).

Tightly regulated factors such as serum retinol and cholesterol are of limited use in cross-sectional validation analyses because individual metabolic differences may dominate between-person variation; however, studies of *changes* in blood levels within subjects may be more informative. In such studies of change, which are more analogous to controlled metabolic studies of dietary effects within subjects, the variability due to strong individual determinants of blood levels is removed, greatly enhancing the opportunity to see smaller effects of diet. Using this approach in a small dietary intervention study, Sacks and colleagues (1986) observed strong correlations between change in saturated fat and linoleic acid intake measured by questionnaire before and during intervention

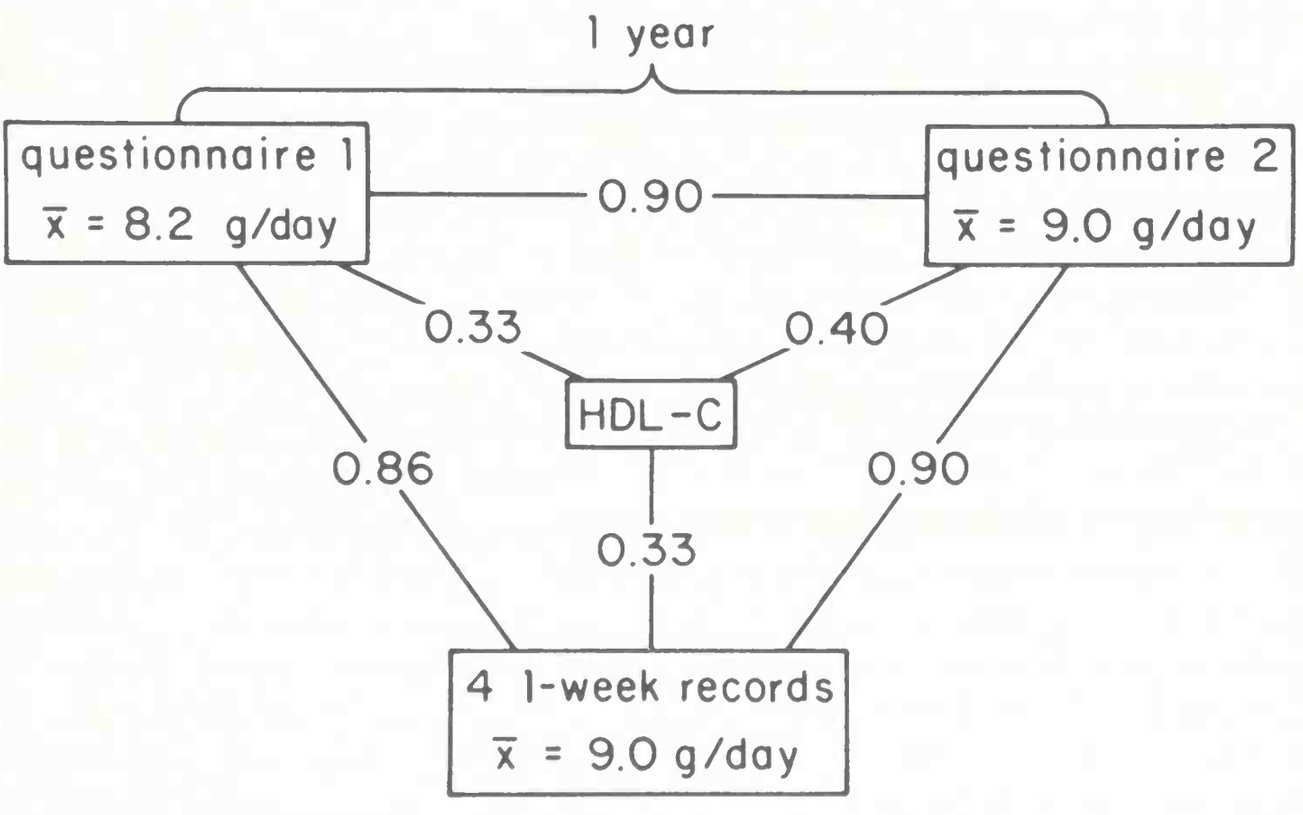


Figure 6–2. Validation of alcohol intake in the Nurses' Health Study. (From Willett et al., 1987; reproduced with permission.)

and change in serum cholesterol. Similarly, Heller and co-workers (1981) found that change in saturated fat intake measured by a simple eight-item questionnaire was associated with change in serum cholesterol over a period of 4 years among participants in an intervention program. In a variation on this approach, we examined the correlation between baseline total vitamin A intake and the change in serum retinol when all subjects were given a daily supplement of 10,000 IU of preformed vitamin A (Willett et al., 1984). The inverse correlation ($r = -0.50$) indicates the anticipated change: subjects with low dietary vitamin A intake had the largest increase in serum retinol when given a fixed supplement.

Although the use of biochemical indicators for purposes of questionnaire validation has been limited, this area deserves further pursuit. As discussed in Chapter 5, it is possible that biochemical indicators may be used to identify maximally discriminating questions or to identify food items that fail to be predictive of biochemical levels despite an appreciable content of the nutrient in question. Such failure to predict nutrient levels could result from an unclearly stated question, incorrect food composition data, or poor bioavailability of nutrient from that food. Resolution of such a finding could lead to improvement in the questionnaire or its food composition database.

Prediction of a Physiologic Response

The ability of a dietary questionnaire to predict an established relationship between a nutrient intake and a physiologic response may be used as qualitative evidence of validity in a manner analogous to using a biochemical indicator. Unfortunately, relatively few such relationships are well established. Because considerable evidence exists that high calcium intake may lower blood pressure, the demonstration of an inverse association between calcium intake and incidence of diagnosed hypertension (relative risk = 0.84, 95% CI 0.71–0.99 for 1000 mg/day compared with less than 400 mg/day) (Witteman et al., 1987) supports the validity of the questionnaire's capacity to measure calcium intake. Sandler and colleagues (1985) have demonstrated that bone density among adult women is positively related to the report of milk consumption many years earlier. This very useful observation not only provides qualitative evidence that the questionnaire can measure milk intake, but also is one of the few elements of data to suggest that recall of remote dietary intake is feasible.

Prediction of Known Disease Relationships

The use of a questionnaire to demonstrate established relationships between a dietary factor and disease can be interpreted as qualitative support for questionnaire validity. This approach is limited by the small number of relationships between diet and disease that are reasonably well established. Examples could include saturated fat intake in relation to coronary heart disease, and green and yellow vegetables in relation to risk of squamous cell lung cancer. As data accumulate and the number of established associations increases, it may become

more feasible to use the demonstration of such a relationship as a measure of validity; at this time this cannot be a primary approach. In a case-control study such associations could well be due to bias; thus it is probably best that such inferences regarding questionnaire validity be based on prospective studies. In a prospective study, however, failure to find an established association can be due to an inappropriate interval of time between dietary assessment and the diagnosis of disease, rather than to low validity of the questionnaire.

DESIGN OF QUESTIONNAIRE VALIDATION STUDIES

Results of studies of diet and disease are often difficult to interpret unless the method used to measure diet has been validated in a population reasonably similar to that being investigated. Without documentation of validity, null associations could simply be due to lack of variation in dietary exposure in the study population or the inability of the dietary assessment method to detect existing differences in diet. The objectives of a validation study can include:

1. Measurement of the true between-subject variation in the dietary factors of interest
2. Qualitative documentation that the epidemiologic assessment method can detect the differences in diet that exist among subjects
3. Standardization of a dietary questionnaire against a true measure of absolute intake, which enhances the capacity to compare findings with other studies
4. Quantitative measurement of exposure misclassification so that measures of association, such as relative risks, can be corrected for measurement error.

The Choice of a Population for a Validation Study

Ideally, the subjects in a validation study should be a random sample of the study population in which the questionnaire is being used. This is often not practically possible, particularly if the population is widely scattered, as face-to-face contact is usually necessary. For example, the Nurses Health Study cohort is national, but logistical reasons necessitated a random sample of Boston-area participants for the validation study (Willett et al., 1985). Based on questionnaires completed by all cohort members, women in the validation study were similar to the other cohort members with respect to intake of most nutrients (unpublished data); however, their use of vitamin supplements was markedly lower than that of women in other parts of the country. To maintain representativeness, maximal efforts to enhance participation are needed once potential validation study subjects are selected, particularly as the time and effort required on their part are substantial. Financial inducements to participation are discouraged by some investigators, but do not seem inappropriate if the time commitment is large, and can be useful if they increase the participation rates.

The Choice of a Comparison Method for a Validation Study

As discussed earlier, diet records (particularly when foods are weighed) usually represent an optimal comparison method as sources of error are largely independent of error associated with a dietary questionnaire. Specifically, the diet record method does not depend on memory, is open-ended, and allows direct measurement of portion sizes. When cooperation or literacy of study subjects is limited, multiple 24-hour recalls may be the best alternative. Although short-term dietary recalls are less demanding than diet recording and are also less likely to influence the actual diet of subjects, their sources of error tend to be more correlated with error in a dietary questionnaire. For some specific dietary factors (e.g., sodium, potassium, beta-carotene, or certain fatty acids), biochemical indicators, such as urinary excretion, blood levels, or adipose tissue concentrations, may be employed (see Chapter 9).

The Choice of an Appropriate Time Frame

Whatever comparison method(s) is selected, it is critical to consider the conceptually relevant time frame. Because "true" intake is usually the average intake over a long period of time, e.g., one or more years, rather than intake over a few days or weeks, it is important that the comparison method reflect this longer time frame. If medium-term variation is likely to exist, such as that due to seasonality, then it is necessary to have multiple measurements during the year. Collection of comparison data over a 1-year period is generally appropriate as seasonal effects and other poorly defined fluctuations in diet are incorporated. Because intake of almost all foods and nutrients varies greatly from day-to-day (see Chapter 3), collection of data from multiple days per subject is essential. Most biochemical markers are also influenced by short-term changes in dietary intake; thus it will also be desirable to obtain replicate measures of biochemical markers, at least for a sample of participants.

The Sequence of Data Collection in a Validation Study

In designing a validation study, the sequence of measurements is of concern as it is possible that the process of collecting one measure may affect response to the other method. In particular, the intensive effort involved in daily recording of diet could conceivably sensitize subjects to their food intake and artificially improve their accuracy in completing a questionnaire administered afterward. On the other hand, administering the questionnaire before the detailed assessment of diet results in an artificially low correlation, as the questionnaire relates to diet before the period of detailed assessment. Because questionnaire administration is relatively inexpensive, the best solution seems to have subjects complete the questionnaire twice, before and after the period of detailed recording. This provides a conservative estimate of the true correlation between the questionnaire and detailed method (provided by the first questionnaire) as well as an optimistic estimate (provided by the repeat questionnaire). The association with

the first questionnaire could also be interpreted as an approximation to the correlation with diet over a period of multiple years.

Because even one year may be a short period in relation to the etiologic effect of diet on cancer and heart disease, it may be useful to extend a validation study to a longer interval. Repeated administration of the questionnaire and collection of biochemical specimens after several years will provide useful data on the longer term stability of diet and biochemical indicators. An additional questionnaire focused on the period of detailed diet data collection several years earlier can provide information on the validity of the questionnaire to measure remote diet, which may be particularly relevant for case-control studies.

Number of Subjects and Replicate Measurements for a Validation Study

Selecting an appropriate number of subjects for a validation study is less than straightforward as correlations for many nutrients are likely to be examined and the precision required (i.e., the confidence intervals around the coefficients) is somewhat arbitrary. Clearly, we always want to know much more than simply whether a correlation coefficient is different from zero.

A realistic degree of desirable precision might be arrived at by considering that correlations for validity generally tend to be in the range of 0.5 to 0.7. If the observed correlation were 0.6, it would be useful to be reasonably sure that it was actually at least 0.4, as levels of validity lower than this will rather seriously attenuate associations (see Chapter 12). The number of participants required to detect this difference in correlations can be calculated based on the standard one-sample formula for sample size [$n = (Z_\alpha + Z_\beta)^2 \sigma^2/d^2$] using Fisher's Z transformation of correlation coefficients (Snedecor and Cochran, 1971). For $\alpha = 0.05$ and $1 - \beta = 0.80$, the number of subjects necessary would be approximately 112. Because one important use of a validation study would be to correct observed relative risks for measurement error (see Chapter 12), another approach for choosing an appropriate sample size would be to consider the implications of various sample sizes on the corrected relative risk estimates and their corrected confidence intervals. Rosner and Willett (unpublished data) have examined the influence of validation study sample sizes on these parameters for different degrees of validity and different observed relative risks. In general, fewer subjects are needed for a validation study with higher degrees of validity. Although no sharp cut-off existed to define an optimal sample size, for realistic conditions (correlations between questionnaire and "truth" of 0.5 to 0.7) it was apparent that validation studies larger than 150 to 200 subjects provide little additional precision in corrected confidence intervals. On the other hand, validation studies with as few as 30 subjects lead to a major increase in the width of corrected confidence intervals. A reasonable size for a validation study thus seems to be about 100 to 200 persons. This is adequate for a range of likely degrees of validity and allows the appropriate deletion of some subjects (e.g., those who become seriously ill or pregnant) during the validation study.

The possibility of using only a small number of replicate measures (e.g., two

24-hour recalls or two 24-hour urines) for the gold standard combined with a statistical adjustment to remove the effects of within-person variation (Rosner and Willett, 1988) is discussed in Chapter 12. As shown in that chapter, similar corrected correlation coefficients were obtained between a food-frequency questionnaire and diet records when either 2 or 4 days of diet recording were used instead of 28 days per subject. In most circumstances, the greatest statistical efficiency (i.e., the lowest variance of the corrected coefficient for a fixed number of measurements) is obtained with only two, and at the most five, replicates per subject. This approach can have major advantages in terms of cost and feasibility compared with the alternative of obtaining large numbers of replicate measurements per subject. The use of corrected correlation coefficients may also provide more valid conclusions in situations where the burden of obtaining large numbers of replicates would eliminate many subjects (and thus reduce the generalizability of the findings), alter the behavior of the participants (and thus the level of the factor being measured), or potentially effect response to another variable (such as a self-administered questionnaire). If a strategy using a small number of replicates per subject is employed, the number of subjects needs to be increased; when the intraclass correlation for replicate measures is low, this may be by a factor of two or three.

DATA ANALYSIS AND PRESENTATION OF VALIDATION STUDIES

Data obtained in a dietary validation study are voluminous as they represent multiple measurements of multiple dietary factors using multiple methods. Reducing this information so as to be useful to the investigator interpreting findings from the main epidemiologic study and to be consumable by readers of the literature presents a formidable challenge.

In analyzing data from a validation study, crude nutrient intakes are of interest, but it is also important to adjust nutrient intakes and biochemical factors for variables that are ultimately controlled in an epidemiologic analysis. For example, age and sex are almost always controlled in epidemiologic analysis so that they should also be controlled in the analysis of the validation study. The reason is that the between-person variation in dietary intake due to these covariates tends to increase observed correlations between methods (as by definition $r = s_b^2/(s_b^2 + s_w^2)$. Because this added between-person variation is removed in the epidemiologic analysis adjusting for such covariates, the artifactually increased correlation coefficients are unrealistic. In other words, for typical epidemiologic purposes it is not useful to know that a questionnaire can detect differences in nutrient intake between young men (which will tend to be high) and older women (which will tend to be low) that are simply due to their age and gender. The effect of controlling for age and sex on correlations between a questionnaire and a 1-year diet record are shown in Table 6–12; correlations decreased appreciably after such adjustment. The rationale for adjusting for total energy intake in epidemiologic analyses is provided in Chapter 11; because it is usually impor-

Table 6–12. Pearson correlation coefficients comparing nutrient intakes from the semiquantitative food-frequency questionnaire with a 1-year diet record

Nutrient	Crude	Calorie-adjusted	Age-, sex-adjusted	Age-, sex-, calorie-adjusted
Total calories	0.67	—	0.37	—
Protein	0.60	0.43	0.46	0.53
Total fat	0.76	0.51	0.57	0.59
Saturated fat	0.74	0.60	0.58	0.62
Linoleic acid	0.74	0.21	0.50	0.28
Total carbohydrate	0.60	0.51	0.37	0.55
Crude fiber	0.44	0.61	0.37	0.65
Cholesterol	0.67	0.38	0.59	0.43
Oleic acid	0.74	0.51	0.55	0.57
Vitamin A	0.63	0.68	0.62	0.70
Niacin	0.48	0.36	0.41	0.37
Vitamin C	0.38	0.46	0.34	0.49
Calcium	0.63	0.55	0.42	0.57
Phosphorus	0.64	0.65	0.52	0.67
Thiamin	0.55	0.41	0.36	0.42
Riboflavin	0.64	0.42	0.55	0.31
Potassium	0.48	0.59	0.33	0.64
Iron	0.47	0.38	0.28	0.40

All variables were transformed by $\log_e$ to improve normality. No supplements are included.
Data were provided by 27 men and women aged 20 to 54.
From Willett et al., 1987.

tant to adjust for total energy intake in case-control or cohort studies, it is important to adjust for energy intake in validation studies as well.

Many alternatives exist for presenting data on the associations between the questionnaire and comparison methods. Although it is useful to compare means and standard deviations for the two methods, it is most important to provide data on the associations between intake measured by the two methods. Alternatives include contingency tables (cross-classification), correlation coefficients, and regression coefficients. Although contingency tables provide a readily interpretable impression of the agreement between two measures (see Table 6–13 for

Table 6–13. Joint classification of calorie-adjusted cholesterol intake assessed by the second semiquantitative food-frequency questionnaire and four 1-week diet records

Questionnaire quintile	Diet record quintile					
	1 (low)	*2*	*3*	*4*	*5 (high)*	*Total*
1 (low)	18	9	2	3	2	34
2	8	11	8	4	4	35
3	4	7	9	9	5	34
4	2	8	9	11	6	36
5 (high)	2	0	7	8	17	34
Total	34	35	35	35	34	173

Data based on information provided by 173 registered nurses aged 34–59 years and residing in the Boston area, 1980–1981.
From Willett et al., 1985.

an example using cholesterol intake), they become cumbersome when many nutrients are being studied. When the relationship between two normally distributed variables is linear (usually reasonably true in the case of validation studies), the correlation coefficient provides, in a single number, essentially the same information as that contained in a contingency table (Walker and Blettner, 1985), thus making this an attractive alternative. Because dietary variables are usually skewed toward higher values, transformations (such as log) to increase normality should be considered before computing correlation coefficients. This has the advantage of reducing the influence of extreme values and of creating a correlation coefficient that can be interpreted in the form of a contingency table. Alternatively, nonparametic correlation coefficients (e.g., Spearman) can be employed when one or both variables are not normally distributed.

A possible disadvantage of the correlation coefficient is that it is a function of the true between-person variation in the population being studied as well as of the accuracy of the questionnaire. This limits the generalizability of the correlation coefficient to those populations with similar between-person variation as the test population. Although this dependence on between-person variation has some disadvantages, this can also be viewed as an advantage as the capacity of a questionnaire to discriminate between subjects, and thus the power of a study, is a function of both its accuracy as well as the between-person variation in exposure within the population.

The relationship between two measures can also be described as a simple regression equation, that is, using the questionnaire (x) to predict true intake (y). This has direct application in correcting relative risk estimates for measurement error, as described in Chapter 12. Lee and colleagues (1983) have advocated the use of regression coefficients for analyzing data from validation studies and state that "two dietary-intake assessments are interchangeable in comparison studies only if they show a linear regression coefficient that is not statistically different from 1.0." It seems, however, unrealistic to expect that any practical method for assessing diet in an epidemiologic study would be so accurate as to be directly interchangeable with a detailed, open-ended method such as multiple weeks of diet recording. Moreover, Wahrendorf (1985) has pointed out that the use of a statistical test is not appropriate for the evaluation of validity; the most certain way to ensure that a regression coefficient is not significantly different from 1.0 would be to conduct a very small validation study.

Because the standard deviations are often different for dietary intake assessed by two methods (i.e., their scales may be different even if they nominally use the same units of measurement), it is usually difficult to assess the capacity of a questionnaire to provide a ranking or relative categorization of subjects on the basis of a regression coefficient. For example, it is possible that a questionnaire provides a perfect ranking of subjects but has either an artifactually large standard deviation (resulting in a regression coefficient less than 1.0) or an artifactually small standard deviation (resulting in a regression coefficient greater than 1.0).

The Kappa statistic, $(P_o - P_e)/(1 - P_e)$, (where P_o is the proportion of subjects observed to be concordant and P_e is the proportion of subjects expected to be concordant on the basis of chance alone) has been increasingly used to com-

pare categories of nutrient intakes measured by two methods. This statistic is an improvement over "percent agreement," as agreement expected on the basis of chance alone is discounted. As pointed out by Maclure and Willett (1987), however, the value of Kappa is not readily interpretable, as it will depend on the number of categories into which a continuous variable has been broken. Thus, its use for comparing ordinal variables, such as nutrient intakes, has considerable disadvantages and should be discouraged.

Bland and Altman (1986) have strongly condemned the use of correlation coefficients for comparing two measures and advocated using the mean and standard deviation of the difference between the two. The standard deviation of the difference is not influenced by the between person variation in exposure; whether this is an advantage or disadvantage has been discussed earlier in the chapter. Interpreting this parameter requires considerable knowledge about the usual absolute values and between-person variation, which differ from nutrient to nutrient. Therefore, the mean and standard deviation of the difference between two methods tends to be cumbersome when evaluating many nutrients, and interpretable only to the very well informed. A related method is to calculate the standard deviation of the residual from the regression of the true measure on the questionnaire measure, sometimes called the *standard error of the estimate*. This residual represents the variation in true intake that is not accounted for by the surrogate measure.

Another presentation of data from a validation study might be called "actual values for surrogate categories." In this approach, subjects are first grouped into categories such as quintiles, on the basis of the surrogate method, in this case the food-frequency questionnaire. Then, the "true value" for these same subjects based on the more detailed method is assigned to the categories defined by the surrogate method. An example of this approach, again based on the validation of the 1980 compressed Nurses Health Study questionnaire, is shown in Table 6–14. The distinct advantage of this method is that it conveys the actual, quantitative differences in diet that correspond to the relative categories defined by the questionnaire. These values are, of course, a function of both the true variation in diet within the population and the measurement error associated with the questionnaire. In Table 6–14 it can be seen that the variation between subjects in total fat intake as measured by the questionnaire is relatively low and that this variation is further reduced by adjustment for calories, even though adjustment for calories increased the correlation coefficient by reducing correlated error. In contrast, the differences between quintiles were much greater for vitamin A in spite of a lower correlation between methods (see Table 6–7) due to the larger true between-person variation in vitamin A intake. These values, of course, need to be established separately for every study population.

These true values for categories of intake assessed by the questionnaire may be useful in presenting epidemiologic data relating to disease risk as they provide simple and direct presentation of the quantitative relationship between level of a dietary factor and risk of disease. In particular, null findings are more readily interpretable as it can be readily determined whether the variation in diet and validity of the questionnaire were adequate to evaluate the relationship being considered.

Table 6–14. Use of actual values for surrogate categories to compare the 1980 Nurses' Health Study dietary questionnaire with 28 days of diet recording by 173 women

	Mean Daily Diet Record Value for Women in Questionnaire Quintile					
Questionnaire quintile	*Total fat (g)*	*Saturated fat (g)*	*Cholesterol (mg)*	*Sucrose (g)*	*Vitamin A[a] (IU)*	*Vitamin C[a] (mg)*
Crude intake						
1	58	20	252	30	4684	69
2	64	24	275	35	5083	117
3	68	23	292	45	6370	151
4	67	26	320	47	7356	218
5	79	28	380	60	8826	436
Calorie-adjusted						
1	61	21	248	32	4259	67
2	64	22	281	39	4761	112
3	67	25	304	42	5795	124
4	70	26	313	50	7099	163
5	71	27	374	51	7593	289

[a]Includes supplements.
From Willett et al., 1985.

This method of data presentation has another practical advantage in that, like regression analysis, it does not require a large number of days of dietary intake per subject to represent "truth." Because the method merely involves computing means for groups defined by the questionnaire, even a single day of diet recording per subject provides unbiased estimates of the actual values for these categories. This method may, of course, be extended to analyses based on specific foods.

Because no single method for relating a surrogate measure to a measure of truth conveys all the available information, it is probably best to present the data in several ways. At a minimum, the means and standard deviations of the true and surrogate measures plus their correlations should be provided. This may be supplemented with other data, such as regression coefficients with standard errors (which may be particularly useful for correcting relative risk estimates) and standard deviations of the residual for true exposure not accounted for by the surrogate measure.

SUMMARY

The interpretation of any study of diet and disease can be substantially enhanced by quantitative information on the validity of the method used to measure dietary intake. Unless the method employed has been previously studied with respect to validity in a similar population, any major dietary study should include a validation component. Because the degree to which data from one validation study can be generalized to other populations is largely unknown at this

time, prudence dictates repeating a validation study if the similarity of populations or circumstances is at all in doubt.

In general, diet records provide the best available comparison method; biochemical markers are potentially useful but do not exist for many dietary factors. Although multiple weeks of diet recording per subject provide the ideal standard, this process is costly. The use of a small number of replicate measures per subject with statistical correction for within-person variation provides an alternative approach that should make a validation study feasible in most epidemiologic settings.

REFERENCES

Abramson, J. H., C. Slome, and C. Kosovsky (1963). Food frequency interview as an epidemiological tool. *Am. J. Pub. Health 53*, 1093–1101.

Acheson, E. D. and R. Doll (1964). Dietary factors in carcinoma of the stomach: A study of 100 cases and 200 controls. *Gut 5*, 126–131.

Balogh, M., J. H. Medalie, H. Smith, and J. J. Groen (1968). The development of a dietary questionnaire for an ischemic heart disease survey. *Isr. J. Med. Sci. 4*, 195–203.

Beaton, G. H., T. Milner, V. McGuire, et al. (1983). Sources of variance in 24-hour recall data: Implication for nutrition and study design and interpretation. Carbohydrate sources, vitamins, and minerals. *Am. J. Clin. Nutr. 37*, 986–995.

Bland, J. M. and D. J. Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet 1*, 307–10.

Bradburn, N. M., L. J. Rips, and S. K. Shevell (1987). Answering autobiographical questions; The impact of memory and inference on surveys. *Science 236*, 157–161.

Browe, J. H., R. M. Gofstein, D. M. Morlley, and M. C. McCarthy (1966). Diet and heart disease study in the Cardiovascular Health Center. *J. Am. Diet Assoc. 48*, 95–100.

Byers, T. E., R. I. Rosenthal, J. R. Marshall, T. F. Rzepka, K. M. Cummings, and S. Graham (1983). Dietary history from the distant past: A methodological study. *Nutr. Cancer 5*, 69–77.

Byers, T., J. Marshall, E. Anthony, R. Fiedler, and M. Zielezny (1987). The reliability of dietary history from the distant past. *Am. J. Epidemiol. 125*, 999–1011.

Colditz, G. A., W. C. Willett, M. J. Stampfer, L. Sampson, B. Rosner, C. H. Hennekens, and F. E. Speizer (1987). The influences of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire. *Int. J. Epidemiol. 16*, 392–398.

Decker, M. D., A. L. Booth, M. Dewey, R. S. Fricker, R. H. Hutchenson, Jr., and W. Schaffner (1986). Validity of food consumption histories in a foodborne outbreak investigation. *Am. J. Epidemiol. 124*, 859–863.

Epstein, L. M., A. Reshef, J. H. Abramson, and O. Bialik (1970). Validity of a short dietary questionnaire. *Isr. J. Med. Sci. 6*, 589–597.

Gordon, T. and D. Shurtleff (1973). The Framingham Study: An epidemiologic investigation of cardiovascular disease. Section 29: Means at each examination and interexamination variation of specified characteristics: Framingham Study Exam 1 to Exam 10. *DHEW Pub. No. (NIH)* 74–478.

Graham, S., A. M. Lilienfield, and J. E. Tidings (1967). Dietary and purgative factors in the epidemiology of gastric cancer. *Cancer 20*, 2224–2234.

Gray, G. E., A. Paganini-Hill, R. K. Ross, and B. E. Henderson (1984). Assessment of

three brief methods of estimation of vitamin A and C intakes for a prospective study of cancer: Comparison with dietary history. *Am. J. Epidemiol. 119,* 581–590.

Hankin, J. H., G. G. Rhoads, G. A. Glober (1975). A dietary method for an epidemiologic study of gastrointestinal cancer. *Am. J. Clin. Nutr. 28,* 1055–61.

Hankin, J. H., A.M.Y. Nomura, J. Lee, T. Hirohata, and L. N. Kolonel (1983). Reproducibility of a dietary history questionnaire in a case-control study of breast cancer. *Am. J. Clin. Nutr. 37,* 981–985.

Heller, R. F., H.D.T. Pedoe, and G. Rose (1981). A simple method of assessing the effect of dietary advice to reduce plasma cholesterol. *Preventive Med. 10,* 364–370.

Hunt, I. F., L. S. Luke, N. J. Murphy, V. A. Clark, and A. H. Coulson (1979). Nutrient estimates for computerized questionnaires vs. 24-hr. recall interviews. *J. Am. Diet Assoc. 74,* 656–659.

Jain, M. G., L. Harrison, G. R. Howe, and A. B. Miller (1982). Evaluation of a self-administered dietary questionnaire for use in a cohort study. *Am. J. Clin. Nutr. 36,* 931–935.

Jensen, O. M., J. Wahrendorf, A. Rosenqvist, et al. (1984). The reliability of questionnaire-derived historical dietary information and temporal stability of food habits in individuals. *Am. J. Epidemiol. 120,* 281–290.

Lee, J., L. M. Kolonel, and J. H. Hankin (1983). On establishing the interchangeability of different dietary-intake assessment methods used in studies of diet and cancer. *Nutr. Cancer 5,* 215–218.

Maclure, M. and W. C. Willett (1987). Misinterpretation and misuse of the Kappa statistic. *Am. J. Epidemiol. 126,* 161–169.

Mullen, B. J., N. J. Krantzler, L. E. Grivetti, H. G. Schutz, and H. L. Meiselman (1984). Validity of a food frequency questionnaire for the determination of individual food intake. *Am. J. Clin. Nutr. 39,* 136–143.

Nomura, A. and J. H. Hankin (1976). The reproducibility of dietary intake data in a prospective study of gastrointestinal cancer. *Am. J. Clin. Nutr. 29,* 1432–1436.

Pietinen P., A. M. Hartman, E. Haapa, et al. (1988a). Reproducibility and validity of dietary assessment instruments: II. A qualitative food frequency questionnaire. *Am. J. Epidemiol. 128,* 667–676.

Pietinen, P., A. M. Hartman, E. Haapa, et al. (1988b). Reproducibility and validity of dietary assessment instruments: I. A self-administered food use questionnaire with a portion size picture booklet. *Am. J. Epidemiol. 128,* 655–666.

Rohan, T. E. and J. D. Potter (1984). Retrospective assessment of dietary intake. *Am. J. Epidemiol. 120,* 876–887.

Rosner, B., C. H. Hennekens, E. H. Kass, and W. E. Miall (1977). Age-specific correlation analysis of longitudinal blood pressure data. *Am. J. Epidemiol. 106,* 306–313.

Rosner, B. and W. C. Willett (1988). Interval estimates for correlations corrected with within-person variation: Implications for study design and hypothesis testing. *Am. J. Epidemiol. 127,* 377–386.

Russell-Briefel, R., A. W. Caggiula, and L. H. Kuller (1985). A comparison of three dietary methods for estimating vitamin A intake. *Am. J. Epidemiol. 122,* 628–636.

Sacks, F. M., G. Handysides, G. E. Marais, B. Rosner, and E. H. Kass (1986). Effects of a low-fat diet on plasma lipoprotein levels. *Arch. Int. Med. 146,* 1573–1577.

Salvini, S., D. J. Hunter, L. Sampson, M. J. Stampfer, G. A. Colditz, B. Rosner, and W. C. Willett (1989). Food-based validation of a dietary questionnaire: The effects of week-to-week variation in food consumption. *Int. J. Epidemiol* (in press).

Sampson, L. (1985). Food frequency questionnaires as a research method. *Clin. Nutr. 4,* 171–178.

Sandler, R. B., C. W. Slemenda, R. E. LaPorte, J. A. Cauley, M. M. Schramm, B. A. Barresi, and A. M. Kriska (1985). Postmenopausal bone density and milk consumption in childhood and adolescence. *Am. J. Clin. Nutr. 42,* 270–274.

Shekelle, R. B., A. M. Shryrock, O. Paul, et al. (1981). Diet, serum, cholesterol, and death from coronary heart disease: The Western Electric Study. *N. Engl. J. Med. 304,* 65–70.

Silverman, D. I., G. J. Reis, F. M. Sacks, T. M. Boucher, M. E. Sipperly, and R. C. Pasternak (1988). Plasma phospholipid EPA levels: A new measure for dietary fish consumption (abstr). *Circulation Suppl. 78,* II-227.

Sokol, R. J., J. E. Heubi, S. T. Iannaccone, K. E. Bove, and W. F. Balistreri (1984). Vitamin E deficiency with normal serum vitamin E concentrations in children with chronic cholestasis. *N. Engl. J. Med. 310,* 1209–1212.

Stefanik, P. A. and M. F. Trulson (1962). Determining the frequency intakes of foods in large group studies. *Am. J. Clin. Nutr. 11,* 335–343.

Snedecor, G. W. and W. G. Cochran (1971). *Statistical Methods.* Ames, Iowa: Iowa State Univ. Press.

Stryker, W. C., L. Sampson, M. J. Stampfer, G. A. Colditz, et al. (1987). Contributions of specific foods to nutrient consumption: Absolute intake vs. between-person variation. Doctoral Thesis, Harvard School of Public Health.

Stryker, W. S., L. Kaplan, E. A. Stein, M. J. Stampfer, A. Sober, and W. C. Willett (1988). The relation of diet, cigarette smoking and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. *Am. J. Epidemiol. 127,* 283–296.

Stuff, J. E., C. Garza, E. O. Smith, B. L. Nichols, and C. M. Montandon (1983). A comparison of dietary methods in nutritional studies. *Am. J. Clin. Nutr. 37,* 300–306.

Thompson, F. E., D. E. Lamphiear, H. L. Metzner, V. M. Hawthorne, and M. S. Oh (1987). Reproducibility of reports of frequency of food use in the Tecumseh Diet Methodology Study. *Am. J. Epidemiol. 125,* 658–671.

United States National Center for Health Statistics (1979). *Dietary Intake Source Data, United States, 1971–74.* Hyattsville, Md.: DHEW Pub. No. (PHS) 79–1221.

Wahrendorf, J. (1985). Re: A comparison of frequency and quantitative methods for epidemiologic studies of diet and disease (letter). *Am. J. Epidemiol. 121,* 776.

Walker, A. M. and M. Blettner (1985). Comparing imperfect measures of exposure. *Am. J. Epidemiol.* 121, 783–790.

Willett, W. C. (1985). Does low vitamin B-6 intake increase the risk of coronary heart disease? In R. D. Reynolds and J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* New York: Alan R. Liss, Inc.

Willett, W. C., M. J. Stampfer, B. A. Underwood, F. E. Speizer, B. Rosner, and C. H. Hennekens (1983). Validation of a dietary questionnaire with plasma carotenoid and alpha-tocopherol levels. *Am. J. Clin. Nutr. 38,* 631–639.

Willett W. C., M. J. Stampfer, B. A. Underwood, L. A. Sampson, C. H. Hennekens, J. C. Wallingford, L. Cooper, C. Hsieh, and F. E. Speizer (1984). Vitamin A supplementation and plasma retinol levels: A randomized trial among women. *J.N.C.I. 73,* 1445–1448.

Willett, W. C., L. Sampson, M. J. Stampfer, et al. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol. 122,* 51–65.

Willett, W. C., R. D. Reynolds, S. Cottrell-Hoehner, L. Sampson, and M. L. Brown (1987). Validation of a semi-quantitative food frequency questionnaire: comparison with a one-year diet record. *J. Am. Diet Assoc. 87,* 43–47.

Willett. W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, and F. E. Speizer (1987a). Dietary fat and risk of breast cancer. *N. Engl. J. Med. 316,* 22–28.

Willett, W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, and F. W. Speizer (1987b). Moderate alcohol consumption and the risk of breast cancer. *N. Engl. J. Med. 316,* 1174–1180.

Willett, W. C., L. Sampson, M. L. Brown, et al. (1988). The use of a self-administered questionnaire to assess diet four years in the past. *Am. J. Epidemiol. 127,* 188–199.

Witteman, J.C.M., W. C. Willett, M. J. Stampfer, G. A. Colditz, F. Sacks, B. Rosner, F. E. Speizer, and C. H. Hennekens (1987). Dietary calcium and magnesium and hypertension: A prospective study. *Circulation (Suppl) 76 (s-4),* iv–35.

7

Recall of Remote Diet

For some diseases, including many cancers, it is hypothesized that the effect of diet may occur many years before diagnosis; thus, the ability to recall diet in the remote past is of considerable interest. Either relatively unstructured diet histories or food-frequency questionnaires can be used to focus questions on a remote period of time.

The validity of remotely recalled diet can be assessed by resurveying individuals with respect to their past diet who had actually provided detailed dietary information in the past, usually for some other purpose. As discussed in the previous chapter, these dietary assessment methods would ideally be different; the original method should be a very detailed assessment, such as diet recording, and the second method one that could be used epidemiologically such as a diet history or food-frequency questionnaire. Even the demonstration that the same method used in the past produces similar data when used to recall remote intake, which is evidence more of reproducibility than validity, can be extremely useful. Several studies of recalled remote diet have been published (Table 7–1).

Jensen and colleagues (1984) found fairly low correlations, ranging from 0.13 to 0.42, between intake originally assessed by a detailed household interview and then recalled by a diet history 15 to 25 years later. These correlations were similar in magnitude to those obtained for frequencies of specific food items at a 20-year interval by Byers and co-workers (1983). Stronger correlations, generally in the range of 0.5 to 0.7 were seen in studies in which the interval between dietary assessments ranged from 3 to 10 years (Rohan and Potter, 1984; van Staveren et al., 1986; Byers et al., 1987; Willett et al., 1988).

The interpretation of remotely recalled dietary data is complex, as diet has some consistency over time. Furthermore, it appears that recalled diet is heavily influenced by current diet. For example, in the studies of Jensen and colleagues (1984), Rohan and Potter (1984), and van Staveren and colleagues (1986), the correlations between recalled past diet and current diet were higher than the correlations between actual past diet and recall of past diet. It is, therefore, possible

Table 7–1. Studies of recalled remote diet

Source	Subjects	Original method	Recall method	Interval	Range of correlation		
					Original vs. recall	*Original vs. current*	*Recall vs. current*
Jensen et al. (1984)	Danish men and women (n = 102)	Household survey	Diet history	15–25 yr	r = 0.13 fresh vegetables to 0.42 energy, and whole milk	r = 0.02 fresh fruit to 0.45 fat, and whole milk	r = 0.32 fish 0.63 energy to 0.73 potatoes
van Staveren et al. (1986)	Dutch men and women (n = 102)	Diet history	Diet history	7 yr	r = 0.46 PUFA to 0.75 total carbohydrate	r = 0.32 PUFA to 0.77 total carbohydrate	r = 0.39 PUFA to 0.75 polysaccharides
Byers et al. (1987)	U.S. men and women (n = 323)	Food-frequency interview	Food-frequency interview	6–10 yr	r = 0.50 fat, 0.61 vitamin A, 0.61 fiber	r = 0.50 fat, 0.49 vitamin A, 0.53 fiber	—
Willett et al. (1988)	U.S. nurses (n = 150)	Four 1-wk frequency records	Food frequency (mailed)	3–4 yr	r = 0.28 iron to 0.61 carbohydrate	—	—
Byers et al. (1983)	U.S. men and women (n = 175)	Food frequency 33 items	Food frequency 33 items	20 yr	r = 0.04 bread to 0.56 coffee	r = 0.00 kale to 0.42 coffee	r = 0.38 bacon, 0.70 head to 0.99 kale
van Leeuwen et al. (1983)	Dutch men and women (n = 79)	7-day record	Diet history	4 yr	r = 0.47 protein to 0.68 fat and 0.82 alcohol	—	—
Rohan and Potter (1984)	Australian men and women (n = 70)	Food-frequency interview	Food frequency (mailed)	3 yr	r = 0.25 protein to 0.78 calcium and 0.87 alcohol	r = 0.35 cholesterol to 0.73 vitamin C and 0.91 alcohol	r = 0.61 PUFA to 0.91 sugar
Bakkum et al. (1988)	Dutch men (n = 40)[a]	Diet history	Diet history	12–14 yr	r = 0.68 carbohydrate to 0.77 energy	r = 0.48 fat to 0.58 energy	r = 0.63 protein to 0.70 fat

[a]24 women also studied but not reported separately.

PUFA = polyunsaturated fatty acids.

that correlations between actual past diet and recalled diet are simply the result of a tendency for diet to remain consistent over time. Until recently, it has been unclear whether the best assessment of remote diet is actually current diet or recall of past diet. These issues have been addressed in several of the previous studies in which three measurements of diet were made: actual past diet, recall of past diet, and current diet.

In examining the assessment of past diet among Dutch adults, van Staveren and co-workers (1986) found similar correlations between either current or recalled diet and the original measurement of past diet. Similar findings were obtained by Jensen and colleagues (1984) in an investigation among a Danish population. Jensen and colleagues also computed partial correlations between recalled and actual diet while controlling for current diet; for each nutrient and food examined this partial correlation was positive, although weakly so, indicating that recall of past diet provided some additional information regarding the original diet beyond that obtained by an assessment of current diet. Byers and colleagues (1983, 1987), however, in substantially larger studies in the U.S., found that actual past diet correlated slightly better with recalled diet than with current diet. Byers and co-workers (1987) also explored the possibility that current diet modified by perceived change in diet would correlate better with the actual past diet; however, recall of past diet remained superior. In a study of over 1000 U.S. men and women who completed food-frequency questionnaires in 1967 and 1982, Thompson and colleagues (1987) found that a retrospective questionnaire completed in 1982 focused on the earlier period provided stronger correlations with the 1967 questionnaire than did an assessment of current diet in 1982. These authors also found that perceived change in the use of food groups reported in 1982 correlated with change measured by the 1967 and 1982 questionnaires (Metzner et al., 1988); however, like Byers and colleagues, they found that the retrospective questionnaire was superior in describing the earlier diet compared with the combination of current diet in 1982 plus the perception of the change. Bakkum and co-workers (1988), in a study among 40 Dutch men, found that a retrospective diet history correlated more strongly with a similar diet history collected 12 to 14 years earlier than did a diet history focused on current intake (Fig. 7–1). Overall, these studies suggest that, if remote diet is of interest, focusing questions on the period of interest provides the most accurate information, although the use of current diet as a surrogate for past diet provides similar information in some instances.

Although remote diet is probably best measured by questions about past diet, the strong influence of current diet on recall raises concerns regarding the possibility of bias when diet has changed after the etiologically relevant period of exposure. This is particularly worrisome in the context of a case-control study when it is possible that the diets of cases, but not controls, may have changed as a result of the diagnosis or the impact of progressive disease and its treatment. This is of increasing concern given the current widespread publicity about diet in relation to cancer and other diseases.

The potential magnitude of recall bias due to the occurrence of disease can be studied by conducting case-control studies within a prospective cohort study of diet and disease. The relationship between dietary factors and disease based

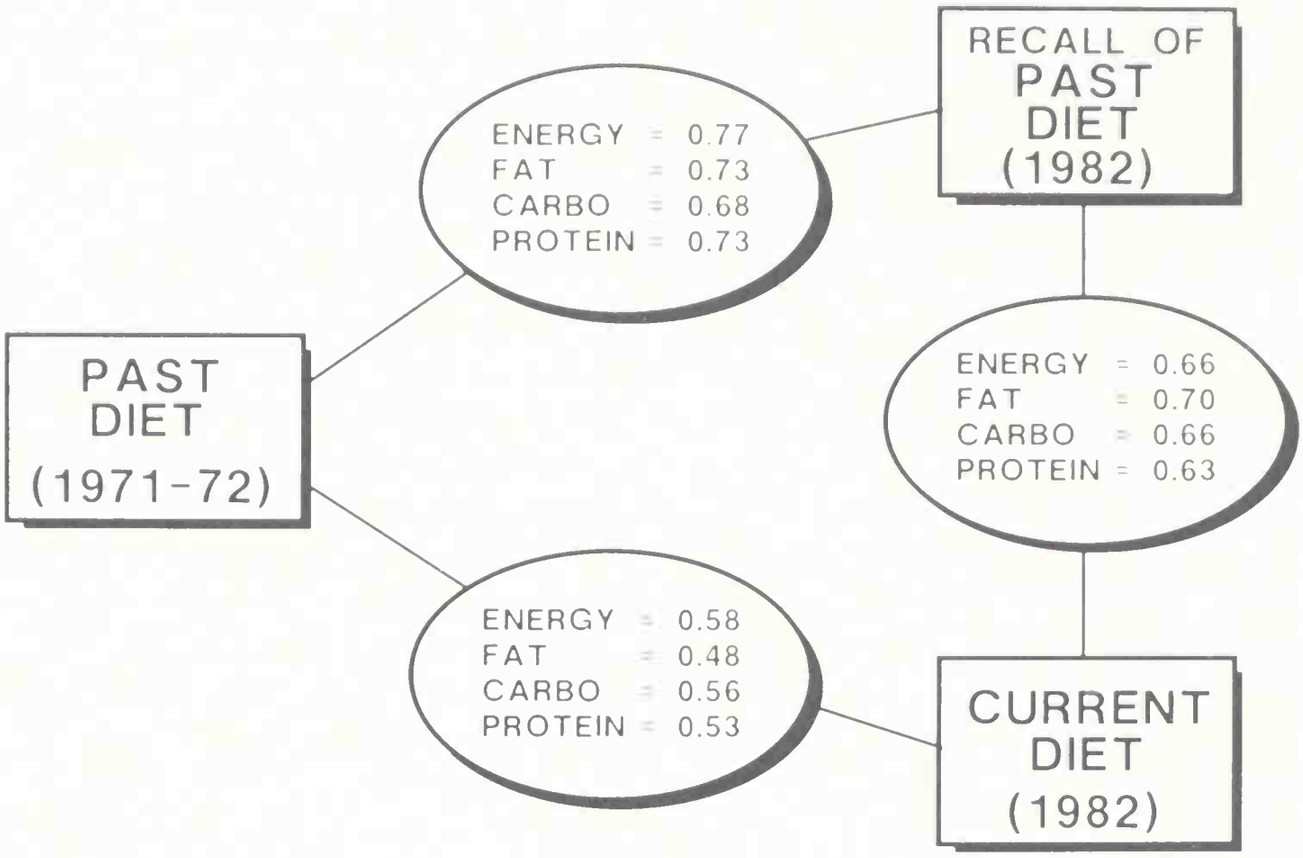


Figure 7–1. Correlations between past diet, recall of past diet, and current diet among 40 Dutch men (data from Bakkum et al., 1988).

on the prospectively collected dietary data would be compared with the same relationship based on dietary data collected after the diagnosis of disease from incident cases and a concurrent sample of subjects without incident disease (the controls).

An assessment of recall bias has been conducted by Bloemberg and colleagues (1986), based on a prospective study of diet and coronary heart disease among 615 men in Zutphen, the Netherlands. The original dietary data were collected in 1970 using a diet history interview, and in 1985 the men were again interviewed regarding their diet 15 years earlier. Analyzed prospectively, the men who subsequently became coronary heart disease (CHD) cases consumed 229 kcal/day less than men who remained free of this disease (a consistent finding in studies of CHD). When analyzed using the retrospective data, however, no difference in energy intake was noted between cases and controls. Although this study suggests that serious recall bias may occur in case-control studies of diet, the time between the diagnosis of CHD and the repeat interview may have been many years, which would rarely be the case in the usual incident series case-control study. Further studies of this general design based on interviews closer to the time of diagnosis are clearly warranted to evaluate the potential for recall bias.

Studies of remote recall have generally used dietary intake data collected in the past for other purposes. There appears to be an unexploited potential for analogous validation studies based on biochemical analyses of specimens collected and preserved for other reasons. For example, adequately stored sera collected in the past could be analyzed for beta-carotene; these levels could then be compared with recalled past intake of carotene-containing foods. Similarly, the

demonstration that recalled history of milk intake during early adulthood (but not recent milk consumption) is correlated with bone density measured many years later provides indirect qualitative validity of recalled calcium intake (Sandler et al., 1985).

Overall, studies conducted to date suggest that diet may be recalled with acceptable levels of misclassification up to approximately 10 years; beyond this period greater uncertainty exists. Validation studies of recalled diet published thus far have only addressed diet during earlier adult years; none have addressed the recall of diet during childhood. Several such studies are ongoing and should provide much-needed data on this topic. Although studies of recalled diet reported to date have focused on the degree of misclassification, future validation studies should also compare alternative strategies for the collection of recalled dietary data so as to refine and improve existing methods. For example, various types of prompts might be compared as well as different levels of detail in the questions about past diet. Further methodologic evaluation and development is required if we are to learn about the long-term health effects of childhood diets by means other than conducting prospective studies that last for many decades.

REFERENCES

Bakkum A., B. Bloemberg, W. A. van Staveren, M. Verschuren, and C. E. West (1988). The relative validity of a retrospective estimate of food consumption based on a current dietary history and a food frequency list. *Nutr. Cancer 11*, 41–53.

Bloemberg, B. P., et al. (1986). The validity of retrospectively assessed dietary intake data in CHD cases and controls (The Zutphen study) (abst). Cardiovascular Disease Newsletter, Council on Epidemiology, American Heart Association, No. 39, p. 52.

Byers, T. E., R. I. Rosenthal, J. I. Marshall, et al. (1983). Dietary history from the distant past: A methodological study. *Nutr. Cancer 5*, 69–77.

Byers, T., J. Marshall, E. Anthony, R. Fielder, and M. Zielezny (1987). The reliability of dietary history from the distant past. *Am. J. Epidemiol. 125*, 999–1011.

Jensen, O. M., J. Wahrendorf, A. Rosenqvist, et al. (1984). The reliability of questionnaire-derived historical dietary information and temporal stability of food habits in individuals. *Am. J. Epidemiol. 120*, 281–290.

Metzner, H. L., F. E. Thompson, D. E. Lamphiear, M. S. Oh, and V. M. Hawthorne (1988). Correspondence between perceptions of change in diet and 15-year change in diet reports in the Tecumseh Methodology Study. *Nutr. Cancer 11*, 61–71.

Rohan, T. E. and J. D. Potter (1984). Retrospective assessment of dietary intake. *Am. J. Epidemiol. 120*, 876–887.

Sandler, R. B., C. W. Slemenda, R. E. LaPorte, J. A. Cauley, M. M. Schramm, M. L. Barresi, and A. M. Kriska (1985). Postmenopausal bone density and milk consumption in childhood and adolescence. *Am. J. Clin. Nutr. 42*, 270–274.

Thompson, F. E., D. E. Lamphiear, H. L. Metzner, V. M. Hawthorne, and M. S. Oh (1987). Reproducibility of reports of frequency of food use in the Tecumseh Diet Methodology Study. *Am. J. Epidemiol. 125*, 658–671.

van Leeuwen, F. E., H. C. deVet, R. B. Hayes, et al. (1983). An assessment of the relative validity of retrospective interviewing for measuring dietary intake. *Am. J. Epidemiol. 118*, 752–758.

van Staveren, W. A., C. E. West, M. D. Hoffmans, P. Bos, A. F. Kardinaal, G. A. van Poppel, et al. (1986). Comparison of contemporaneous and retrospective estimates of food consumption made by a dietary history method. *Am. J. Epidemiol. 123,* 884–893.

Willett, W. C., L. Sampson, M. L. Browne, M. J. Stampfer, B. Rosner, C. H. Hennekens, and F. E. Speizer (1988). The use of a self-administered questionnaire to assess diet four years in the past. *Am. J. Epidemiol. 127,* 188–199.

8

Surrogate Sources of Dietary Information

JONATHAN M. SAMET

Many diseases for which dietary hypotheses have been advanced are either immediately or rapidly fatal. For example, median survival after the diagnosis of lung cancer is only 5 months; the prognosis of stomach cancer is equally poor (Axtell et al., 1976). In studying such diseases with either a case-control or a retrospective cohort study design, surrogate sources of information on diet may be requisite because the index subject is deceased or too ill to provide dietary information. Furthermore, in the setting of severe illness, subjects may alter their dietary pattern so that current consumption, as assessed by a diet record or a recall method, does not reflect the intake relevant to the hypothesis under study. Reporting of past consumption by ill subjects may also be affected by illness-related changes in present consumption. Such changes may also alter levels of biochemical markers assessed when the subject is ill.

Although subjects who cannot provide dietary information directly can be excluded from an etiologic study, bias results if diet affects prognosis. For example, if dietary consumption of beta-carotene were a risk factor for small cell lung cancer but not for other types, exclusion of deceased cases would introduce bias because small cell lung cancer has the poorest prognosis. Exclusion of deceased and ill cases may also lengthen the duration of a study and may be impracticable for some diseases such as fatal myocardial infarction.

Instead of excluding deceased and ill subjects from studies of diet and disease, dietary information may be obtained by interview with a surviving spouse, sibling, child, or other informant. The spouse's diet may also be used as a surrogate measure (Kolonel and Lee, 1981). Complete dietary information, however, may be unavailable from surrogate sources and differential and nondifferential misclassification may result from their use. To the extent that spouses share the same dietary pattern, biochemical measurements made on samples from a healthy spouse might be appropriate surrogates for the ill spouse.

In some settings, it is possible that surrogate responses may be less biased than those provided by an index subject. For example, in the context of a case-

control study of an illness that alters the diet of the cases, surrogate responses might provide a more accurate assessment of past diet. Although random misclassification might be increased by such use of surrogate respondents, differential misclassification might be reduced. This potential advantage of using surrogate responses has not been evaluated, but merits consideration.

This chapter reviews the limited literature on surrogate sources of dietary information. The availability of dietary information from surrogates, the validity of such information, and the analysis of data from surrogates are considered.

AVAILABILITY OF INFORMATION FROM SURROGATE RESPONDENTS

Pickle and colleagues (1983) described the availability of information from surrogate respondents in three case-control studies. Although the questionnaires did not include diet, their findings for cigarette smoking have implications for dietary studies. For all cigarette smokers, questions were asked on the numbers of cigarettes smoked daily and on the duration of each level of daily smoking. These questions are comparable with those needed to establish temporal variation of consumption of a particular food. Complete detailed information on smoking could be supplied by only 56 percent of spouses, 47 percent of siblings, and 55 percent of offspring. The overall pattern of response suggested that siblings were the best information source for information about a subject's immediate family or events that occurred during early life, whereas spouses and children were the best information source for events during adult life.

In a population-based case-control study of lung cancer in New Mexico, interviews were conducted with surrogates for approximately half of the cases (Samet et al., 1985). Dietary data were collected with a food frequency approach designed to assess consumption of vitamin A, both the preformed vitamin A and the carotenoid precursors. For each food, the subjects were asked their usual frequency of intake, as measured on a nine-level scale, during a 1-year reference period. For the lung cancer cases, the reference period for those with eating habits unaffected by their disease was the year ending on the interview data. For cases with a change in eating habits that preceded diagnosis, the reference year ended at the time of the change. For control subjects, the reference year was chosen based on the average patient delay interval for lung cancer patients in New Mexico. Usual portion size was determined by reference to pictures of food servings, measuring devices, and convenient units, such as one egg.

Responses for 36 foods were combined to create indices of total vitamin A, preformed vitamin A, and carotene consumption. The extent of missing data for these foods is provided in Table 8–1. For the control and the self-reported case interviews, complete information was obtained for nearly all subjects. In contrast, interview with a surrogate for the case yielded complete information for about 80 percent of the cases. For approximately 10 percent, the indices of vitamin A could not be calculated because of the extent of the missing data.

Lerchen and Samet (1986) interviewed 80 wives from 1983 through 1984 for the same histories provided earlier by their husbands, who were cases in the lung

Table 8–1. Availability of dietary information from controls, index cases, and case surrogates in a case-control study of lung cancer

	Controls		Cases			
			Self-report		*Surrogate-report*	
Missing items	Frequency (%)	Amount (%)	Frequency (%)	Amount (%)	Frequency (%)	Amount (%)
None	97.9	96.1	96.8	94.8	83.1	81.7
1–4	1.8	3.7	1.7	3.5	5.9	7.1
5–9	0.0	0.0	0.0	0.3	0.9	1.5
10–14	0.0	0.3	0.0	0.0	0.0	1.2
≥15	0.3	0.0	1.5	1.4	10.1	8.5

Based on responses to a food-frequency questionnaire. For frequency, 47 items were considered in tabulating the missing data; for amount, 36 items were considered. Amount, usual portion size, considered to be unavailable if frequency not reported. Information for all controls was self-reported. The data were obtained from the case-control study reported by Samet et al., 1985.

cancer case-control study in New Mexico. The dietary component of the questionnaire was limited; the wives were asked to report their husbands' usual frequency of intake of six foods selected from the original questionnaire: red and green chile, carrots, liver, eggs, and peaches. All of the wives could provide this information for their deceased husbands.

The results of these studies demonstrate that surrogate respondents can provide dietary information, but that incomplete responses must be anticipated (Table 8–1). The findings of Pickle and co-workers (1983) suggest that the availability of dietary information varies with the relationship of the surrogate respondent to the index subject.

COMPARABILITY OF INFORMATION FROM SURROGATE RESPONDENTS

The comparability of dietary information from a surrogate respondent with that given by an index respondent has been assessed in studies involving simultaneous interviews with spouse pairs. These investigations have not directly addressed the validity of surrogate responses; none has included comparisons of the interview data supplied by the respondents and their surrogates with other measures of intake, such as diet records. The principal studies of spouse pairs are further limited by not replicating the usual circumstances of interview with a surviving partner in a case-control or a retrospective cohort study.

The findings from four recent studies of the comparability of self-reported and spouse-reported dietary information (see Tables 8-2 to 8-6) are reviewed in this chapter. Moore and colleagues (1970) also performed a relevant study but a small sample size limits interpretation of their findings.

Kolonel and co-workers (1977) interviewed 300 pairs of adult subjects concerning the eating habits of the husband. The subjects were participants in a health survey in Hawaii and were all residents of urban areas of the island of Oahu. Few of the surrogate respondents were elderly; 56 percent were younger

than 45 years of age and only 5 percent were 65 years or older. A questionnaire on dietary habits, alcohol consumption, and smoking was administered to both members of the pair by an interviewer. The spouses were interviewed separately and not permitted to discuss the questions together.

Weekly frequency of consumption was determined for 13 foods and for beverages (Table 8–2). The mean frequencies of consumption as reported by the index subjects and by their wives were very close. The level of agreement between the two interviews varied with the frequency with which a food was consumed; it tended to be lower for more frequently consumed foods and higher for less frequently consumed foods. The extent of agreement for the foods was comparable with that for age started to smoke and for the number of cigarettes smoked.

A similar study was reported by Marshall and colleagues (1980) (Table 8–3). The index subjects included 67 male cases and 91 male controls in a case-control study in New York state. From 2 weeks to 4 months after the interview with the index subject, the spouse completed the same interviewer-administered questionnaire. Frequency of consumption of an extensive list of foods was determined with an 11-category scale. For assessment of agreement, this scale was reduced to four categories.

The findings (Table 8–3) closely paralleled those reported by Kolonel and colleagues (1977). Mean monthly frequencies were similar as reported by the men and by their wives for them, although the extent of agreement did not vary consistently with the frequency of consumption. The extent of agreement within one unit tended to be extremely high. The limited range of a four-level scale and nonuniform distribution among these categories, however, would insure a high level of agreement within one unit. The investigators did note, however, that extreme disagreements were infrequent; in less than 2 percent of the spouse pairs were responses given that placed the frequency in both the maximum and the minimum categories.

In another study, 46 subject spouse pairs were interviewed with a food-fre-

Table 8–2. Comparability of food consumption as reported by male respondents and their spouses for them in a study in Hawaii

	Mean weekly frequency		Agreement	
Food	*Index subject*	*Spouse*	*Exact*	*± 1 Unit*
Processed meats	3.0	2.9	62.7 (%)	79.0 (%)
Beef	4.3	3.9	55.9	75.6
Pork	1.0	1.0	88.6	96.3
Poultry	1.5	1.5	85.0	96.3
Eggs	2.9	2.7	72.2	88.0
Rice	6.3	6.9	65.0	75.7
Pickled vegetables	2.3	2.3	70.2	78.9
Raw vegetables	5.7	5.7	50.7	64.7
Fresh fruit	5.2	4.8	52.0	67.7
Milk	0.7	0.7	66.6	94.0

Abstracted with permission from Tables 2 and 3 in Kolonel et al., 1977.

Table 8–3. Comparability of food consumption as reported by male respondents and their spouses for them in a study in New York State

Food	Mean weekly frequency		Agreement	
	Index subject	*Spouse*	*Exact*	*± 1 Unit*
Bacon	5.6	5.3	57.0 (%)	90.1 (%)
Hamburger	6.4	6.3	74.7	97.4
Liver	1.1	1.2	61.0	97.4
Poultry	3.8	4.1	59.3	96.0
Cheese	7.0	7.6	56.2	85.0
Milk	16.3	17.3	59.4	83.2
Broccoli	1.5	1.7	66.7	95.5
Carrots	4.7	5.2	55.6	91.5
Lettuce	12.1	12.4	61.7	92.9
Tomatoes	10.9	11.3	53.7	91.3
Summer squash	1.2	1.3	62.9	87.4
Apples	9.6	10.2	52.8	87.4
Melons	5.4	6.0	52.6	83.6
Peaches	7.6	9.2	45.2	84.9

Ascertained with 11 categories but collapsed to four for assessment of agreement.
Abstracted with permission from Tables 1 and 2 in Marshall et al., 1980.

quency questionnaire designed to provide an index of vitamin A consumption (Humble et al., 1984). In 38 pairs, each person answered questions both as a subject and as a surrogate for his or her spouse. An additional eight subject surrogate pairs came from eight couples with each member serving either as a subject or as a surrogate only. A 55-item food-frequency questionnaire, which ascertained frequency of use on a nine-level scale, was administered with reference to the year preceding the date of interview. The size of a usual serving was assessed on a six-level scale with food models of standard portions. The past pattern of consumption was categorized as more, the same, or less than the pattern during the reference year.

For individual foods, the findings for frequency of consumption (Tables 8–4 and 8–5) were similar to those reported by Kolonel and colleagues (1977) and by Marshall and colleagues (1980). For portion size, the level of agreement varied from 29 to 100 percent and was 64 percent when averaged across all foods. The average agreement for past pattern of consumption was 57 percent.

Humble and colleagues (1984) used the responses concerning frequency and amount of consumption to calculate two indices of total daily vitamin A intake: one based on frequency alone and the other based on both frequency and amount. For men, similar levels of mean daily consumption were obtained from their own responses and those of their wives. In contrast, the husbands' responses yielded lower estimates of total daily vitamin A consumption for their wives than were calculated from the wives' own responses. For example, for the index based on frequency and portion size, the mean based on the women's responses was 8602 units, whereas that based on the husbands' responses was 6518 units ($p < 0.05$).

Agreement of index subject-based and spouse-based indices of total con-

Table 8–4. Comparability of food consumption as reported by male respondents and their spouses for them in a study in New Mexico

	Mean weekly frequency		Agreement	
Food	*Index subject*	*Wife*	*Exact*	*± 1 Unit*
Liver	0.0	0.0	93 (%)	100 (%)
Cheese	0.48	0.55	60	95
Whole milk	0.21	0.20	76	95
Eggs	0.38	0.37	67	100
Broccoli	0.19	0.19	57	100
Carrots	0.27	0.26	48	100
Lettuce	0.76	0.81	45	97
Tomatoes	0.78	0.60	51	95
Summer squash	0.13	0.10	52	93
Cantaloupe	0.16	0.20	62	100
Peaches	0.31	0.25	19	90

Ascertained on a nine-level scale.

Abstracted with permission from Table 2 in Humble et al., 1984.

sumption was assessed by Spearman rank order correlation and by comparison of quantile groupings. The extent of agreement tended to be higher for the index based on frequency alone (Spearman r = 0.49 for men and r = 0.52 for women). Concordance of the quantile rankings ranged from 26 to 52 percent and varied with the type of index and the choice of quantiles: quintiles, quartiles, and tertiles. For example, with the frequency-based index, the concordance of quintiles was 31 percent for men and 38 percent for women.

Herrmann (1985) has reported the most recent investigation on the comparability of information from next-of-kin respondents. The subjects were male and

Table 8–5. Comparability of food consumption as reported by female respondents and their spouses for them in a study in New Mexico

	Mean weekly frequency		Agreement	
Food	*Index subject*	*Husband*	*Exact*	*± 1 Unit*
Liver	0.01	0.01	86 (%)	100 (%)
Cheese	0.62	0.54	43	93
Whole milk	0.21	0.26	81	88
Eggs	0.33	0.34	69	100
Broccoli	0.16	0.20	74	98
Carrots	0.33	0.24	40	93
Lettuce	0.86	0.92	53	97
Tomatoes	0.80	0.75	55	93
Summer squash	0.14	0.14	52	88
Cantaloupe	0.27	0.29	57	93
Peaches	0.39	0.31	43	93

Ascertained on a nine-level scale.

Abstracted from Table 3 in Humble et al., 1984.

female cases and controls in a colon cancer case-control study in Pennsylvania. The index respondents all completed an interviewer-administered food-frequency questionnaire that was referenced to a calendar year several years earlier. The next-of-kin were assigned at random to complete either the interviewer-administered questionnaire or a self-administered questionnaire. The frequency questions were open-ended on the former but were categorized on the latter.

Exact agreement was calculated for cases and controls by the type of interview with the next-of-kin (Table 8–6). The agreement was lower for the self-completed questionnaire, but was comparable for men and women. Herrmann (1985) also noted that agreement tended to be better for the spouse than for other next-of-kin respondents.

Lerchen and Samet (1986) described agreement between the responses of index cases and those of their surviving spouses for six foods. As in the other investigations, the means based on the two sets of responses were similar. Exact agreement on a five-level scale was 45 percent for carrots, 32 percent for peaches, 51 percent for eggs, 52 percent for liver, 48 percent for green chili pepper sauce, and 41 percent for red chili pepper sauce. Spearman rank order correlation coefficients ranged from 0.38 to 0.56.

The results of these studies (see Tables 8–2 to 8–5) demonstrate close agreement between mean intakes of specific foods reported by interview subjects and their surrogate next-of-kin. The comparability of the mean intakes implies that surrogate data may be useful for descriptive studies of diet and disease; the surrogate responses appear to provide a minimally biased estimate of mean consumption by a group. Thus, surrogate reports on food consumption can be used in ecological studies of diet and disease, such as cross-cultural or international comparisons.

Table 8–6. Comparability of frequency information from cases and controls and from their next-of-kin in a case-control study of colon cancer in Philadelphia

	Exact agreement by interview type	
Food	Interview (%)	Self-completed (%)
Cases and Next-of-Kin		
Beverages	77	65
Vegetables	70	49
Fruits	66	49
Meats	71	59
Dairy	67	49
Grains	75	58
Controls and Next-of-Kin		
Beverages	77	69
Vegetables	66	53
Fruits	63	52
Meats	63	62
Dairy	63	53
Grains	70	63

Abstracted with permission from Table 6 in Herrmann, 1985.

Surrogate responses may be less useful for case-control and cohort studies. The extent of agreement for individual foods is variable (Tables 8–2 to 8–6) and depends on the frequency with which a food is consumed, the sex of the index subject, the relationship of the surrogate to the index subject, and the method of data collection. The analyses of Humble and colleagues (1983) demonstrates that agreement between subject-based and surrogate-based measures of overall consumption may be less satisfactory than had been suggested by previous studies based on individual foods.

The studies reviewed are limited in the extent to which they replicate the usual circumstances leading to interview with a surrogate respondent. Kolonel and colleagues (1977) and Humble and colleagues (1984) recruited healthy volunteers and both studies did not include a large proportion of elderly subjects. Only the small investigation reported by Lerchen and Samet (1986) included spouses of deceased index subjects.

USE OF SPOUSE DIET AS A SURROGATE

In several epidemiologic studies of diet and cancer, the diet of one spouse has been used as a surrogate for the diet of the other. For example, Nomura and coworkers (1978) conducted a breast cancer study in wives of men in the Honolulu Heart Study. The diet of the men whose wives had breast cancer was compared with that of the remaining cohort members. In another example, the occurrence of large bowel cancer was assessed in surviving spouses of persons who had died from cancer of the colon or of the rectum (Jensen et al., 1980). Underlying the investigation was the assumption that married couples have a similar diet for a long period of time.

The diets of husbands and wives, however, are not strongly correlated. Kolonel and Lee (1981) described the correlation between the diets of 281 spouse pairs in Hawaii. Average frequency of food consumption tended to be comparable for the men and the women. Exact agreement on frequency was less than 50 percent for most foods and intraclass correlation coefficients ranged from 0.33 to 0.55 for food items, from 0.15 to 0.32 for nonalcohol containing beverages, and from −0.07 to 0.66 for alcoholic beverages. Lee and Kolonel (1982) subsequently reported an expanded investigation that included 1428 couples. Nutrient indices were derived from reported intake of 83 food and beverage items. On average, men consumed more than women, but intake per unit body weight was similar. Spearman and intraclass correlation coefficients showed moderate concordance between the intakes of the spouses. The Spearman correlation coefficients for the nine-nutrient indices considered in the report ranged from 0.45 to 0.59; the intraclass coefficients ranged from 0.35 to 0.59.

The findings of these studies suggest that the diet of a spouse may serve as a surrogate when another source of information is unavailable. They do not establish, however, that the spouse's diet is a preferable surrogate to interviewing the surviving case or control subject. Comparisons of associations obtained from a typical case-control study in which subjects themselves are interviewed with

those from interviews of spouses of the same cases and controls would be of considerable methodologic interest.

ANALYSIS OF DATA FROM SURROGATE RESPONDENTS

Although the extent and direction of bias from surrogate respondents remain unclear (Gordis, 1982), the possibility of bias must be considered in any data based on surrogate responses. The potential for bias appears to be greatest when data for one group, such as cases, are obtained from surrogates, whereas data for another, such as controls, are obtained from the index subjects. As discussed by Gordis (1982), the direction of information bias has not been documented and might plausibly be in either a positive or a negative direction. As suggested by several of the studies reviewed previously, substantial nondifferential misclassification must also be anticipated. Thus, even if unbiased mean values for cases or controls are provided by surrogate sources, loss of precision in estimating individual intake may substantially bias relative risks and other measures of association toward null values.

In using data from surrogate respondents, certain analytic strategies should be followed. First, the extent of missing data for the various respondent groups should be assessed and a determination made concerning the extent of missing information that is acceptable. For example, in analyzing data from a lung cancer case-control study, Samet and colleagues (1985) excluded subjects with missing data on either frequency or portion size for four or more foods. For those with missing items for three or fewer foods, the modal responses for the appropriate respondent group were substituted. Second, the responses of the surrogate and index subjects should be compared within homogenous groups for which comparable responses would be anticipated.

Finally, in testing associations in a case-control or cohort study, respondent type must be considered as a potential confounding or modifying variable. Analytic strategies include stratification by respondent type, or introduction of indicator variables for respondent type into multivariate models (Samet et al., 1985). The demonstration of substantial heterogeneity by respondent type, even without a formally significant test of interaction, precludes analyses of the complete data.

CONCLUSIONS

Reliance on surrogate sources of information may be unavoidable in investigating rapidly fatal illnesses. The literature, however, indicates that next-of-kin information concerning diet may introduce misclassification in analytic investigations of dietary hypotheses. Both differential and nondifferential misclassification must be considered. For dietary hypotheses, where the effects of interest are often either unknown or anticipated to be relatively small, misclassification from any source may lead to erroneous conclusions. Use of surrogate sources of dietary information should be minimized in studies of diet and disease.

REFERENCES

Axtell, L. M., A. J. Asire, and M. H. Myers (1976). *Cancer Patient Survival.* Report number 5. U.S. Department of Health, Education, and Welfare, Bethesda, Md.: DHEW Publication No. (NIH) 77-992.

Gordis, L. (1982). Should dead cases be matched to dead controls? *Am. J. Epidemiol. 115,* 1–5.

Herrmann, N. (1985). Retrospective information from questionnaire. I. Comparability of primary respondents and their next-of-kin. *Am. J. Epidemiol. 121,* 937–947.

Humble, C. G., J. M. Samet, and B. E. Skipper (1984). Comparison of self- and surrogate-reported dietary information. *Am. J. Epidemiol. 119,* 86–98.

Jensen, O. M., A. M. Bolander, P. Sigtryggsson, M. Vercelli, X. Nguyen-Dinh, and R. MacLennan (1980). Large-bowel cancer in married couples in Sweden. A follow-up study. *Lancet 1,* 1161–1163.

Kolonel, L. N., T. Hirohata, and A. M. Nomura (1977). Adequacy of survey data collected from substitute respondents. *Am. J. Epidemiol. 106,* 476–484.

Kolonel, L. N. and J. Lee (1981). Husband–wife correspondence in smoking, drinking and dietary habits. *Am. J. Clin. Nutr. 34,* 99–104.

Kolonel, L. N., J. H. Hankin, J. Lee, S. Y. Chu, A. M. Nomura, and M. W. Hinds (1981). Nutrient intakes in relation to cancer incidence in Hawaii. *Br. J. Cancer 44,* 332–339.

Lee, J. and L. N. Kolonel (1982). Nutrient intakes of husbands and wives: Implications for epidemiologic research. *Am. J. Epidemiol. 115,* 515–525.

Lerchen, M. L. and J. M. Samet (1986). An assessment of the validity of questionnaire responses provided by a surviving spouse. *Am. J. Epidemiol. 123,* 481–489.

Marshall, J., R. Priore, B. Haughey, T. Rzepka, and S. Graham (1980). Spouse–subject interviews and the reliability of diet studies. *Am. J. Epidemiol. 112,* 675–683.

Moore, M. C., E. M. Moore, C. D. Beasley (1970). Dietary–artherosclerosis study on deceased persons. *J. Am. Diet Assoc. 56,* 23–28.

Nomura, A., B. E. Henderson, and J. Lee (1978). Breast cancer and diet among the Japanese in Hawaii. *Am. J. Clin. Nutr. 31,* 2020–2025.

Pickle, L. W., L. M. Brown, and W. J. Blot (1983). Information available from surrogate respondents in case-control interview studies. *Am. J. Epidemiol. 118,* 99–108.

Samet, J. M., B. J. Skipper, C. G. Humble, and D. R. Pathak (1985). Lung cancer risk and vitamin A consumption in New Mexico. *Am. Rev. Respir. Dis. 131,* 198–202.

9

Biochemical Indicators of Dietary Intake

DAVID HUNTER

This chapter considers the utility of biochemical indicators to represent the dietary intake of specific nutrients in epidemiologic studies. This use of biochemical indicators is different from their common roles as measures of nutrient status or as predictors of disease risk. Nutritionists and clinicians frequently use a biochemical indicator as a measure of a person's current ability to meet physiologic requirements for that nutrient. Nutrient intake is just one determinant of nutrient status, however, because the levels of a nutrient in blood or tissues can be affected by genetic and lifestyle factors such as smoking or physical activity, or the intake of other nutrients. A related use of biochemical indicators is to predict disease risk, irrespective of whether the level of the biochemical measure is determined by dietary intake or other factors. For example, serum cholesterol strongly predicts risk of cardiovascular disease, although the association of serum cholesterol with dietary cholesterol is weak, due to the influence of nondietary factors such as genetic variation, exercise, and obesity on serum levels (see Chapter 1). As nutritional epidemiologists, we have a primary interest in the intake of nutrients as quantifiable determinants of disease, and thus, our use of a biochemical indicator is principally as a measure of nutrient intake. Thus, our intent in this chapter is to review the relationships between the intake of individual nutrients and biochemical indicators, rather than to discuss the use of biochemical measurements to predict nutrient status or disease.

First, we discuss some general principles of the relation between nutrient intake and biochemical indicators, which define the utility of these measures for epidemiologic purposes. We then review some aspects of study design, specimen collection and storage, and measurement relevant to the validity of biochemical indicators. We next discuss the study designs available for validation of biochemical indicators as measures of intake. To illustrate these principles we use, as an example, selenium, a nutrient for which the relation between dietary intake and levels in various tissues is unusually well characterized. Finally, we review

specific biochemical indicators available to assess intake of selected nutrients of current epidemiologic interest.

There are two fundamental uses for biochemical indicators of food and nutrient intake in epidemiologic studies. The most common application is as a surrogate for actual dietary intake in studies of disease occurrence. For nutrients that vary widely in concentration within individual foods and for which food composition tables are inaccurate, a biochemical measure may be the only feasible way to estimate intake. Within-food variation may occur because of differences in food storage, processing, or preparation, or may be due to geographic differences in soil nutrient content. For example, Ullrey (1981) demonstrated that the selenium content of corn within the United States can vary by as much as 100-fold, and that, in turn, the selenium content of swine muscle varies more than 15-fold. A less common use for biochemical markers, but one that is likely to receive more attention in the future, is to validate other forms of diet assessment. In other words, if a biochemical indicator provides a good index of true intake, then the performance of other assessment techniques can be assessed by comparison with it.

GENERAL PRINCIPLES

Biochemical measures are objective and may, therefore, have a superficial aura of impartiality, especially to the epidemiologist trained to be suspicious of the answers provided by forgetful and potentially biased human subjects and their interviewers. It is important, however, to be aware that these measures are almost always subject to the same problems of misclassification and bias. The following general principles concern the issues of error in classification with respect to intake, and the avoidance or control of sources of bias.

Sensitivity to Intake

The most important requirement of any index of food or nutrient intake is, of course, that it measures what it is supposed to measure. At the extreme, if there is no correlation between blood level of a nutrient and true intake, then the blood levels randomly classify subjects with respect to intake. The magnitude of this correlation depends principally on the degree of homeostatic control of the levels of the nutrient in available biologic samples, on the range of intake to be investigated, and on the existence of determinants other than intake.

Fortunately for humans (if not for nutritional epidemiologists) homeostatic mechanisms control the concentration of many nutrients in body tissues and fluids. These mechanisms may be conceptually simple, such as the saturation of absorption (e.g., iron) or the excretion of excess (e.g., vitamin C), or they may involve complex hormonal pathways (e.g., calcium). These mechanisms mean that the relation between nutrient intake and levels in biologic specimens is rarely linear, and may not even be monotonic. For many nutrients, these mechanisms cause the increase in biologic levels to be attenuated, or plateau, with higher nutrient intake (Fig. 9–1).

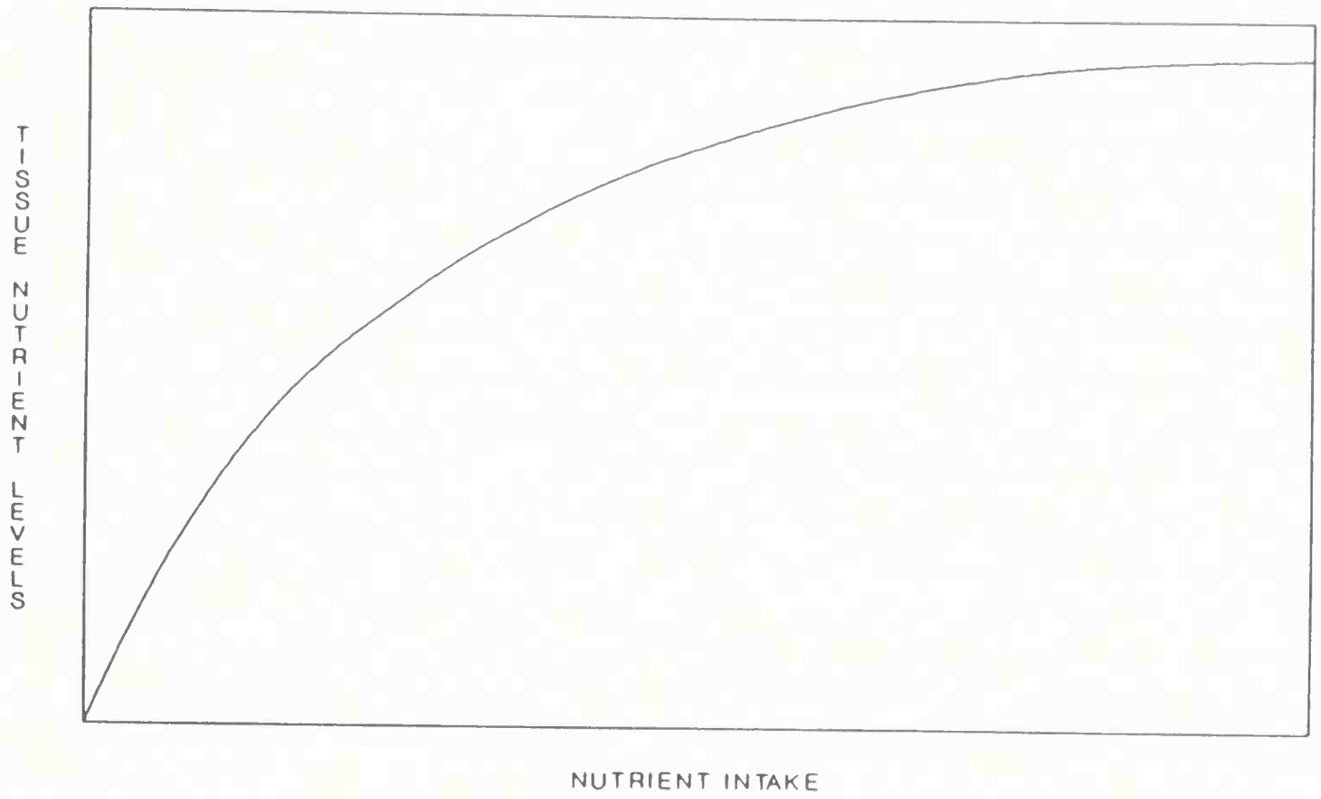


Figure 9–1. Typical relation between nutrient intake and tissue nutrient levels.

The implications of these relationships for biochemical indicators of nutrient intake are several. If the plateau phase of a marker is wide, its concentration may be almost uniform over the range of normal intake, and thus the indicator is almost useless as a marker of nutrient intake. This is also the case if the range of intake of the study population is relatively narrow and falls in the plateau phase. If the population is more heterogeneous with respect to intake, then the biochemical measure may, at least, be able to discriminate between very low and very high intakes.

Another relevant consideration in the context of sensitivity to intake is the question of bioavailability. Bioavailability has been defined as "the proportion of a nutrient in food that is absorbed and utilized" (O'Dell, 1984). Some chemical forms of a nutrient may be absorbed far less completely than others. Thus, a biochemical indicator may accurately reflect the intake of a well-absorbed nutrient species, but fail to reflect the intake of a slightly modified compound. This may be a disadvantage if we are interested in total dietary intake of the nutrient, such as in hypotheses involving nutrient actions within the intestinal lumen. More often, however, we are concerned with the biologic effects of nutrients, after intestinal absorption. In this regard, biochemical measurements may be more appropriate than estimates of intake based on food composition tables, as these rarely contain information on the relative bioavailability of nutrients.

Time Integration

Having stressed that the first requirement for a valid biochemical indicator is sensitivity to intake, we must now point out that (in this context) it is possible to have too much of a good thing. Nutrient exposures relevant to disease, like most exposures in chronic disease epidemiology, are usually long-term. The

induction period of atherosclerotic plaques or the promotion period for cancers, may be years or decades. Thus, it is usually desirable that a biochemical indicator reflect the cumulative effect of diet over an extended period of time (Willett, 1987). If the indicator is only sensitive to short-term intake, and if nutrient intake fluctuates from day-to-day (as it does for most nutrients), then the biochemical indicator reflects nutrient intake for the hours, days, or weeks before sampling, not the months, years, or decades that are truly of interest.

Time-integration is largely a function of the nature of the biologic sample being tested. "Longer-lived" specimens are more likely to reflect long-term intake. Nutrients stored in fat tissue, for instance, usually have a slow turnover. For many elements, concentrations of nutrients in erythrocytes are less susceptible to short-term fluctuations in concentration than in plasma or serum. A 24-hour urine collection is more likely to be representative of intake than a random urine sample. Tissue levels, such as liver retinol concentration, may be the best indicators of long-term nutrient status; however, these are rarely obtainable in epidemiologic studies. Hair and nails are readily available keratinous tissues, which take weeks or months to grow, and thus offer the prospect of measuring long-term intake.

Information about time-integration comes from several sources. Experimental studies, in which nutrient intakes are manipulated, provide good information on the relation between nutrient intake and the time-course of response in different biologic specimens. Less direct but still valuable information can be obtained by simply sampling levels in individuals longitudinally. A high within-person variance in a biochemical measure over time implies a response to short-term fluctuations in intake, or to other determinants of its level. In either case, the power of a single measurement to predict long-term average level is low if the within-person variation is large. Sensitivity to recent intake can be overcome to a certain extent by repeated sampling, and Liu and colleagues (1979a) have given a method to calculate the number of replicate measures required to estimate the "true" underlying mean value within a specified range of error (also see Chapter 3). Even if a biochemical measure is a good indicator of dietary intake over the previous months or years, repeated sampling may still be desirable to account for individual change and secular trends in nutrient intake.

Measurement of Nondietary Determinants of Biochemical Indicator Levels

Dietary intake and homeostatic mechanisms are usually not the only determinants of nutrient concentrations in most biologic samples. A wide array of genetic, environmental, and life-style factors may also influence nutrient levels. If these factors are not taken into account, then misclassification of subjects with respect to nutrient intake will occur. If such a factor is not associated with disease, then it contributes to random misclassification and attenuates any real association between the biochemical indicator and disease. If the factor *is* associated with disease, then confounding may occur. For example, vitamin E levels in plasma or serum are considerably influenced by the concentration of cholesterol in the blood as this lipid-soluble vitamin is transported by the lipoproteins.

Among premature infants, low serum vitamin E was not found to be related to clinical vitamin E deficiency unless adjusted for the high lipid levels associated with the liver dysfunction of prematurity (Sokol et al., 1984). The observation that a biochemical indicator of dietary intake is correlated with another factor does not automatically mean that marker levels should be adjusted for this factor; legitimate question may exist as to whether an adjustment should be used. An understanding of the biologic basis for a correlation between the biochemical indicator, such as plasma vitamin E, and the other factor, such as plasma cholesterol, helps resolve this issue. If plasma vitamin E levels and plasma cholesterol were correlated due to an influence of vitamin E intake on serum cholesterol, adjustment would not be appropriate as controlling for an effect of the primary exposure, whether or not that effect is related to the biologic outcome, would represent "overcontrol." In this example, however, we have information from a randomized trial indicating that vitamin E intake does not affect serum cholesterol levels (Willett et al., 1983a). If, on the other hand, plasma cholesterol levels influence plasma vitamin E levels, as appears to be the case, it may still be unclear whether the variation in plasma vitamin E due to difference in plasma cholesterol is biologically relevant to vitamin concentrations in the tissues of the target organ. The observation among premature infants noted previously, that blood vitamin E levels were not related to clinical vitamin E deficiency unless adjusted for serum cholesterol, provides direct evidence that variation in vitamin E levels due to blood lipids is not biologically relevant and that adjustment is appropriate. Thus, information on whether an association with disease is strengthened by adjustment of the biochemical indicator for another factor can be extremely useful. In addition, information on whether the association between nutrient intake and the biochemical indicator is increased by adjustment for the other factor provides direct guidance as to whether the adjustment is appropriate. Thus, the finding that the correlation between calculated vitamin E intake and plasma vitamin E increased from $r = 0.11$ ($p = 0.19$) to $r = 0.28$ ($p = 0.02$) after adjustment for plasma cholesterol (Willett et al., 1983b) suggests that if blood vitamin E levels are to be used as an indicator of dietary intake, they should be adjusted for blood cholesterol and probably other lipid fractions as well.

The same strategies used to control confounding or reduce random error in other contexts are appropriate here. Subjects may be matched on characteristics ascertainable in advance by the investigator, or restriction rules may be applied. Information on potential confounders may be collected for control in subsequent analyses. Many biochemical measurements vary with age and sex; however, these two variables are almost always routinely controlled for in epidemiologic analyses. As in many other fields of epidemiology, knowledge about other potential confounders is incomplete, and the possibility of uncontrolled confounding must be considered in any assessment of biochemical data.

Types of Analytic Procedures

There are several broad approaches to biochemical indicator measurement. In a sense, these represent different views about the relevant aspects of biochemical

assessment, that is, whether tissue nutrient levels or functional status of the subject is most important.

Direct Measurement

The most obvious method is direct assay of the concentration of a nutrient, or its metabolic products, in a tissue or a fluid. As new and more refined methods have become available that reduce limits of detection and increase accuracy, direct measurement of more nutrients in a wider variety of specimens has become feasible. As discussed previously, the usefulness of a potential indicator of intake has to be determined individually for each nutrient, and for each type of biologic specimen.

Functional Assays

An intuitively appealing alternative to the measurement of nutrient concentrations is the measurement of biochemical functions that depend on specific nutrients. The activities of several nutrient-specific enzymes (e.g., glutathione peroxidase (selenium); erythrocyte transketolase (thiamin)) have been shown to reflect dietary intake of these nutrients. Care must be taken, however, both that the system used is truly nutrient-specific, and that its function has not been compromised by some aspect of specimen collection or storage.

In the broadest sense, tests that measure ultimate products of metabolic pathways are tests of the adequacy of intake of participating nutrients (e.g., iron and hemoglobin). Few, if any, however, of these pathways are nutrient-specific, as nonnutrient causes of abnormalities usually exist. Nonbiochemical tests of physiologic function (e.g., vitamin A and the dark adaptation test) may provide an index of nutrient inadequacy, but rarely provide a measure relevant to the full range of normal intake.

Tolerance Tests

Tolerance tests usually involve taking a blood or urine sample shortly before, and some time after, the administration of a test dose of the relevant nutrient (e.g., vitamin A and the relative dose response test) (Campos et al., 1987). Poorly nourished individuals frequently retain a larger proportion of the test dose than well-nourished persons, who may simply excrete most of the test dose. The increased logistic demands of these methods limit their use in many epidemiologic studies.

STUDY DESIGN CONSIDERATIONS

Specimen Collection and Storage

Appropriate timing of sampling and proper collection and storage techniques are essential to ensure the validity of biochemical measures. Ideally, all steps involved in the process of delivering biologic specimens to the waiting analyst would be designed such that the concentration of the indicator to be measured is exactly the same as it is in vivo. Biologic reality and the complexities and

time-course of epidemiologic studies mean that this ideal is rarely attainable. A first requirement of validity, therefore, is that case and control specimens must be handled in the same way. This must be done to avoid introduction of systematic differences due to study design. A systematic difference, for instance, in the time of storage between the case and the control specimens, could bias a study result if the biochemical measure is degraded over time. Substantial degradation in storage of both case and control specimens would, of course, eliminate the possibility of finding a difference, even if a difference existed before specimen collection. The storage requirements for different nutrients vary considerably, and, for many, information on stability is sparse.

Seasonal Timing

Nutrient intakes may vary in response to the availability of seasonal foods. For many nutrient markers, therefore, it is important that samples from both cases and controls be collected in the same season, preferably in the same year. If this is not feasible, then adjustment should be made for season of collection in the statistical analysis of the data.

Time of Day

Nutrient markers that are responsive to recent intake may vary substantially over the course of a single day. The collection of early morning, fasting blood samples, or overnight urine specimens, usually minimizes this source of variation. Even measures of long-term intake may be influenced by diurnal variation, and standardization of the time of collection is thus desirable. Again, if this is not possible, time of collection should be controlled for in the analysis.

Options in Study Design

In case-control studies, it may already be too late to collect the relevant samples from cases if the onset, diagnosis, progression, or treatment of the illness affect the levels of the indicator of interest. Results derived from biochemical indices used in case-control studies have not always been confirmed in prospective studies, often because the disease was found to influence nutrient levels (see Chapter 13). One strategy for dealing with this problem is to attempt to enroll cases as soon as possible after the date of diagnosis, or even while they are undergoing a diagnostic work-up. Many diseases, and especially many cancers, however, have extended preclinical periods, and changes in dietary patterns, or the secretion of tumor products, may affect nutrient levels months before the cases have been identified.

Even prospective studies may be prone to this form of bias. For example, low serum cholesterol was associated with increased cancer risk in the 2 years subsequent to blood collection in the MRFIT study, however, this effect attenuated over time, and may well have been due to preclinical disease (Sherwin et al., 1987). Care should thus be taken in the interpretation of biochemical results in the first years of follow-up. In any case, the long induction periods probably

associated with many chronic diseases mean that dietary intake years, or decades, before disease onset is the exposure of interest.

The opportunity for efficient and cost-effective assessment of exposures years before disease onset offered by the *nested* case-control design deserves consideration. In a nested case-control study, cases of a disease arising in a cohort are compared with a sample of other cohort members, rather than to the entire cohort. If the exposure of interest requires measurement of a biochemical parameter in a stored specimen, this design permits a statistical analysis based on a small fraction of the biochemical analyses that would be required if the relevant biochemical parameter was measured for all members of the cohort. Willett and colleagues (1983c), for example, had available to them frozen sera from 4480 men and women who had participated in a large clinical trial of treatment for hypertension, as well as information on cancer incidence in the 5 years subsequent to blood collection. Their exposures of interest were retinol, vitamin E, and selenium, and a conventional cohort study approach would have required them to obtain laboratory analyses for measurements on all 4480 subjects. Using a nested case-control design, however, they matched two control subjects to each of the 111 cancer cases, and evaluated the effects of these exposures using data on 333 individuals instead of 4480. This approach also preserves a high proportion of specimens for subsequent analyses related to other endpoints, leaving the biologic specimen "bank" relatively intact. Nested case-control studies are an efficient method for maximizing the utility of a large cohort study.

Another design option for maximizing the efficiency of the use of biologic specimens is the possibility of pooling specimens. Aliquots of samples from individuals with a common characteristic (e.g., cases or controls) may be combined and analyzed as a single unit (Peto, 1983). This method provides an obvious economy by replacing many measurements with a single one, an advantage if the analytic test is particularly expensive. In addition, if the test requires a relatively large minimum sample volume relative to the total amount of sample available for each individual, the administrators of the specimens may be very reluctant to lose a large proportion of their samples. Pooling permits a smaller aliquot from each individual to be contributed, and thus may be particularly suited for screening hypotheses for more labor-intensive and specimen-intensive investigation. Wahrendorf and colleagues (1986) measured retinol, beta-carotene, and alpha-tocopherol in 610 blood samples, and found good relative agreement for retinol and beta-carotene between values measured in pools of 50 specimens, and the means of the 50 individual values contributing to each pool. The major disadvantage of the pooling approach is that information is lost on the range of nutrient levels in each comparison group, and the mean levels are susceptible to outlying values about which only individual measurements yield information. In addition, the capacity to address issues of confounding and interaction is reduced. Because only a small number of laboratory measurements are made, the potential influence of technical error becomes large and needs to be carefully considered. Indeed, the expected differences between cases and controls are typically similar to the coefficient of variation for many laboratory measurements.

Contamination

Contamination may occur before, during, or after specimen collection. Precollection contamination is essentially restricted to environmentally exposed biologic specimens such as hair or nails. Hair may be particularly prone to superficial contamination due to its relative porosity and high surface-to-volume ratio. Selenium-containing shampoos, for instance, probably invalidate the use of hair as a substrate for selenium measurement in countries in which their use is widely prevalent (Morris et al., 1983). Attempts to deal with surface contamination by washing must be evaluated on a nutrient-by-nutrient basis, as the possibility of leaching relevant nutrients out of the sample must be considered.

Contamination during, or after, collection can occur for all types of specimens. The definition of what constitutes contamination relates to the original concentration of the nutrient in question (Iyengar, 1987). Nutrients, such as trace elements, with extremely low concentrations in biologic samples, are more susceptible to a small amount of introduced extraneous material than nutrients with a greater original concentration. Much of the published data on levels of many trace elements in biologic samples may be erroneous due to the lack of appreciation of the potential for contamination (Versieck, 1985). All needles, syringes, containers, and solutions that the specimen comes into contact with should be examined for their potential to contaminate the sample. The common anticoagulant EDTA, for instance, contains measurable quantities of iron, zinc, chromium, selenium, and other trace elements (Iyengar, 1987). Standard blood collection tubes and rubber stoppers may be contaminated with zinc; thus, specially prepared tubes should be used for zinc estimation (Smith et al., 1985). Indeed, the declining trends in reported normal values of several trace elements over recent decades may reflect greater control of contamination, rather than any underlying secular change in intake (Iyengar, 1987). The articles by Versieck (1985) and Iyengar (1987) provide comprehensive reviews of current knowledge about contamination and trace element analysis.

Stability

Few nutrient measurements relevant to epidemiologic studies can be performed immediately after sampling. Some storage time is thus usual, and specimens may be stored for years or even decades in nested case-control studies. The stability of the relevant nutrients is thus a matter for concern.

If there is any possibility that a nutrient is not completely stable during storage, then identical handling of case and control specimens is again a first requirement for preventing bias. This may not always be an easy requirement to fulfill. If cases are hospital-derived, for instance, and controls are sought in the community, then exact matching of the time of day of specimen collection, temperature control, and the length of time between collection and analysis, may be difficult to achieve. A log should be kept for each specimen of all events such as sample handling and freezer failure, so that the history of sample storage is known to investigators at the time of analysis. For many nutrients, information is limited on the effects of different methods of long-term storage. This poses

something of a problem for the epidemiologist designing a prospective study, or contemplating the analysis of a nutrient in an existing bank of biologic material. In general, simple elements may be less prone to degradation than complex proteins, and samples in which metabolic activity had ceased at the time of sampling, such as hair and nails, may be more stable than those in which enzymatic activities are potentially ongoing. The most common approach to long-term storage of blood components and urine has been freezing in an ordinary freezer (−20°C) or an ultra-low freezer (approximately −70°C). Lyophilization (freeze-drying) may have advantages for certain samples.

Specimen Quality Controls

Quality control procedures are as essential in the monitoring of collection and storage procedures, as they are in the monitoring of the actual chemical analysis. Quality control, in this context, consists of two activities. Initially, the protocol for collection and storage must be validated with respect to protection from contamination. These procedures should be periodically reviewed. Nutrient stability during storage must also be assessed. This may be accomplished by forming a pool of samples and storing aliquots for periodic analysis. As the aliquots should have the same initial nutrient levels, monitoring these levels over time should provide information about sample stability and "laboratory drift." This procedure also allows calculation of the true between-run coefficient of variation percent (CV% = [SD/mean] × 100%), relevant to the time-course of an epidemiologic study in which samples may be analyzed over a period of months or years.

ANALYSIS

Practicing epidemiologists rarely have the time, the resources, or the laboratory expertise, to actually perform the procedures involved in nutrient analysis. It is important, therefore, to maintain a close liaison with laboratory colleagues and to be aware of potential problems that will affect the interpretation of the exposure measure. Epidemiologists may obtain a rapid initial evaluation of the utility of a laboratory test by sending the laboratory blind specimens that reflect the population range of exposure to the relevant nutrient, and splitting each specimen into two or more aliquots. Calculation of the between-person CV% gives useful information on the ability of the test to discriminate between individuals, and an analysis of variance comparing different aliquots from the same person provides an assessment of the extent of random laboratory error.

Overview of Analytic Techniques

Much of the stimulus to use biochemical indicators has been derived from the advances that have been made in the measurement of the relevant compounds in biologic tissues. New methods have decreased the random error of measurements so that the small differences in means that are usually relevant for epi-

demiologic studies (see Chapter 3) may be detectable. Improvements in the lower limit of detection for many nutrients, particularly the trace elements, has meant that new exposure measures are now available.

Minimizing Systematic Error Bias

Reducing random measurement error augments the power of a study to detect between-group differences in that measurement. Reduction of the variance due to random error is not only the domain of the analytic chemist, but may depend on the amount of money the epidemiologist (or funding agency) is prepared to spend on expensive analytic methods. Minimization of systematic misclassification or observation bias may be largely achieved by blinding the laboratory to case-control status, and by eliminating systematic differences in the way case and control specimens are handled. Laboratory drift, or between-assay variation, occurs in most analytic techniques. It is thus desirable to analyze all specimens in a single run, if possible. Paired cases and controls should be analyzed consecutively (with the within-pair order randomized), and pairs should be randomly ordered with respect to variables of interest, so that effects of order (within-run laboratory drift) are not attributed to another variable. If it is not possible to analyze all samples in a single run, then ensuring that cases and their controls are analyzed together at least ensures the validity of the comparison.

Analytic Quality Control

Quality control is an essential aspect of any program of laboratory analysis. Analytic quality assurance has been called "an attitude of mind" (Parr, 1985), implying that constant awareness of the possibilities of error may be as important as adherence to formal procedures in the pursuit of analytic validity. Regular quantitative estimation of error, both random and systematic, is important. Controlling random error is largely a matter of estimating precision, and fine-tuning analytic methods to improve it. The usual measure of precision is the CV%. The within-run CV% is determined by dividing single samples into two or more aliquots and analyzing them together. To estimate the between-run CV%, aliquots are analyzed in different runs, usually on different days. The larger the CV% inherent in a measurement technique, then the larger will be the sample size necessary to derive stable estimates of the group mean. In general, one expects the within-run CV% to be less than the between-run CV%, as there is more potential between runs for variation in room temperature, buffer concentrations, personnel, and so on. The between-run CV% is largely irrelevant if all samples are included in a single run. It is difficult to generalize about acceptable numerical values for the laboratory CV%, as the degree of error acceptable depends on the number of samples available, the mean concentration of the nutrient, between-person variation, and a number of other factors. Reported values of within-run CV% for modern nutrient analytic procedures are often 5 percent or less, and this is well within the range of random error tolerable in most epidemiologic studies. It is important to note, however, that published values may represent the performance of techniques in the hands of highly experienced

investigators under optimal conditions, and that more error may attend samples run in large batches, in other laboratories, or at later dates.

The other principal component of a quality assurance program is the use of quality control standards, either internal standards prepared within the laboratory, or certified reference materials prepared by other laboratories. This provides not only a test of within-laboratory procedures, but allows comparison with other laboratories using the same certified materials, permitting more appropriate interpretation, or even pooling, of the results of different studies. Ideally, reference materials should be used to validate the analytic procedure over the whole relevant working range, not just a single point.

VALIDATION OF BIOCHEMICAL MEASURES AS INDICATORS OF DIETARY INTAKE

Before a biochemical indicator may be used as a measure of nutritional intake, it must be evaluated with respect to its sensitivity to that intake. A validation procedure should also assess, if possible, the other important determinants of indicator concentration, so that adjustment can be made for these in study design or analysis. Epidemiologists whose principal concern is the relation between diet and disease may be somewhat less interested in exploring the effects of diet on tissue nutrient levels, explorations that seem more appropriate for the metabolic enthusiast. If these indicators are to be used as measures of dietary exposure, however, then the epidemiologist is obviously responsible for ensuring that the exposure measure is a valid representation of long-term intake. In this respect, epidemiologists' requirements of biochemical indicators may be different from those of many nutritionists, who frequently prefer short-term measures of response to therapy. Several strategies are available to define the relation between long-term dietary intake and biologic levels.

Animal Studies

The exposure of laboratory animals to nutrients can be tightly controlled, and validation of a biochemical indicator can be attempted by measuring it in animals fed different nutrient intakes. Salbe and Levander (1987), for instance, demonstrated that feeding selenium to rats resulted in increases in hair and nail selenium, in a dose-response manner. This approach, however, suffers from the general problem of animal studies; that is, concern about the generalizability of the conclusions to humans.

Geographic Correlation of Intake and Biologic Markers

Specimens from areas of known deficiency of a specific nutrient can be compared with specimens from average and high exposure areas. Thus, Morris and co-workers (1983) correlated geographic differences in selenium intake with group differences in nail selenium concentration. For example, specimens from New Zealand, an area known to have low selenium intake, were uniformly lower

than any in the United States (Fig. 9–2). With respect to blood measures, studies in New Zealand comparing residents with visitors, indicate that plasma selenium concentration is more sensitive to short-term changes than erythrocyte selenium, which provides a selenium index integrated over several months, presumably due to the relatively slow turnover of red blood cells (Rea et al., 1979; see Fig. 9–4). This method has the appeal of using human subjects directly, under natural circumstances, without the potential ethical problems inherent in dietary manipulation. The major problem is that associated with all geographic correlation studies: the difficulty of identifying and controlling confounding factors. International differences in serum cholesterol, for instance, may be due to differences in exercise, body composition, or alcohol consumption, as well as differences in cholesterol intake. An important advantage of this approach is that it provides the opportunity for the exposure measure to reflect the long-term intake of settled individuals. Long-term nutrient intake is of greatest interest for most epidemiologic purposes, and also the hardest to simulate in dietary manipulation studies.

Correlation with Individual Intake

An observational approach to validation based on individuals is to estimate their dietary exposure and determine its relation with biochemical levels. One efficient way to do this is to select individuals on the basis of known exposure (e.g., vitamin supplement users). Thus, among U.S. women consuming selenium supplements, a linear dose-response relationship was observed between supplement dose and toenail selenium level (Hunter et al., 1987). Alternatively, members of the general population provide information over the range of usual population exposure. Another efficient approach if the relevant nutrient has already been measured in a large number of samples, is to identify individuals with extreme values and attempt to assess their dietary intake of the nutrient. Dietary intake may be estimated from the analysis of duplicate meals, from diet records, recalls, or food-frequency questionnaires. As long-term intake is usually the exposure of interest, methods with a long reference period are preferable. If the biochemical indicator has been independently validated and is known to reflect intake, then the study design can be reversed, and diet assessment methods can be validated against the biochemical marker.

Dietary Manipulation in Humans

A rigorous test of the relation between nutrient intake and a biochemical indicator may be achieved if a dietary manipulation study is feasible. These trials are usually limited to weeks or months, however, so the measure of dietary intake may be relatively short-term. These studies can take advantage of many of the desirable features of clinical trials, such as randomization, placebo-control, or crossover designs. If the intervention involves extensive restructuring of normal diet, however, blinding of the subjects to treatment may not be possible. Studies of vitamin and mineral supplementation, for instance, have been invaluable in delineating the relation between intake and biochemical indicators,

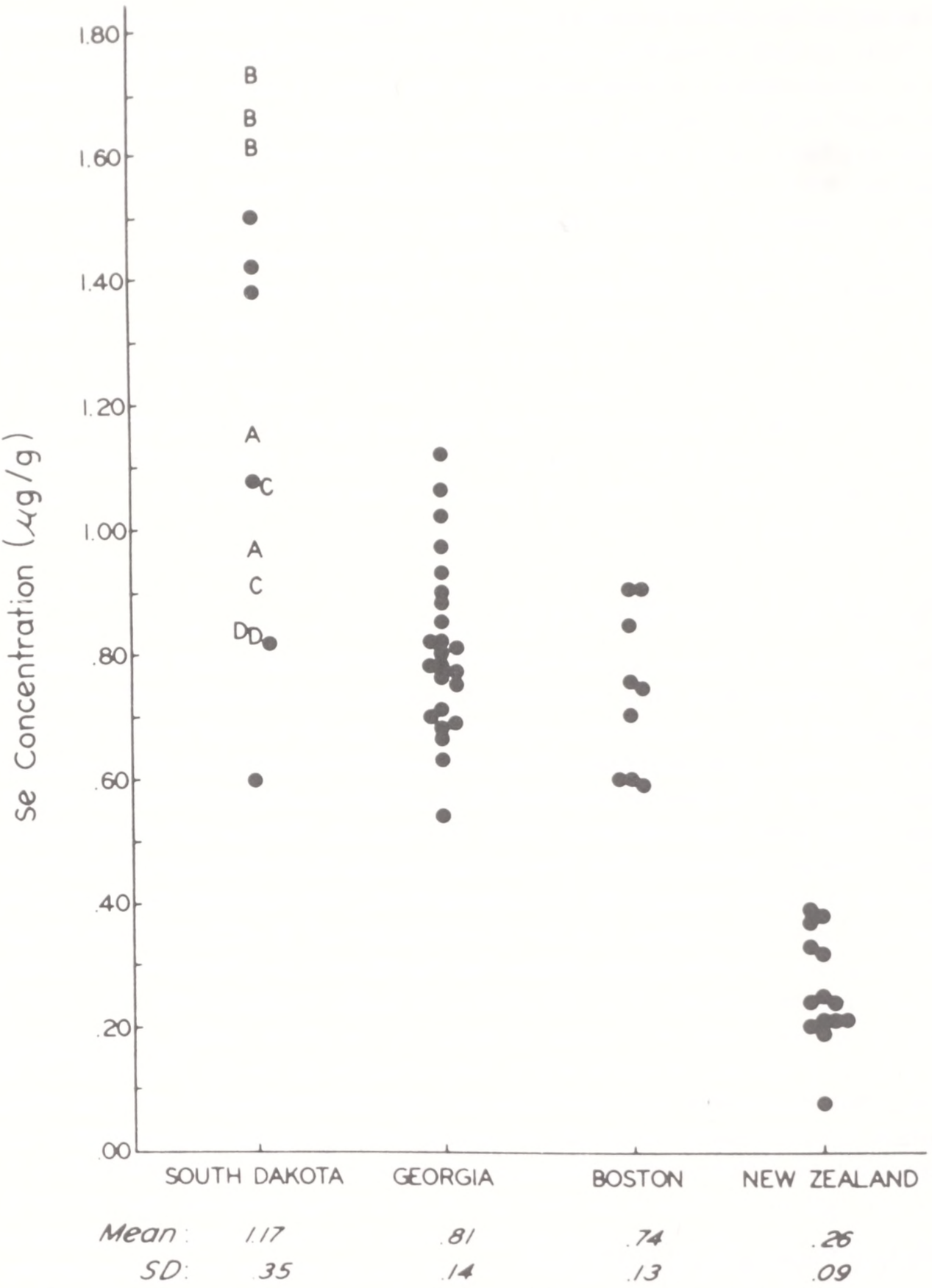


Figure 9–2. Selenium concentration in toenails from four regions. Selenium exposure is known to be unusually low in New Zealand, intermediate in Boston and Georgia, and unusually high in South Dakota. Letters indicate members of the same family, dots indicate unrelated persons. (From Morris et al., 1983; reproduced with permission.)

allowing, in some instances, complete assessment of the time-integration of the dose-response relations. Most studies designed to evaluate indicators for epidemiologic use are performed on free-living subjects. If a more complete understanding of nutrient pathways is desired, then metabolic ward conditions may be necessary. It is desirable to collect information on other determinants of the biochemical measure to control for these in the analysis.

Several experimental studies of selenium intake have provided valuable information on the use of various biologic tissues as indices of selenium exposure. Selenium supplementation increases serum, plasma, and erythrocyte selenium (Thomson et al., 1985; Luoma et al., 1985; Neve et al., 1988). In an experiment to determine the metabolic fate of an oral dose of isotopically labeled selenomethionine, erythrocyte selenium concentration increased and washed out more slowly than plasma selenium, implying a lower sensitivity of erythrocyte selenium to transient changes in exposure (Griffiths et al., 1976). Among 10 Belgian adults supplemented for 60 days with 100 μg Se/day as selenomethionine, plasma selenium was significantly increased after only 5 days of treatment, whereas erythrocyte selenium increased only after 45 days (Neve et al., 1988). Similarly, when four New Zealand women consumed 200 μg/day of selenium in a bread made from selenium-rich wheat, plasma and whole blood selenium levels increased earlier and decreased sooner in the postdosing phase, than erythrocyte levels (Thomson et al., 1985). In an earlier experiment, erythrocyte selenium was still increasing after over 300 days of a selenomethionine supplement (Robinson and Thomson, 1981) (Fig. 9–3). Erythrocyte selenium, therefore, appears to be the preferable blood indicator of long-term exposure status. Whole

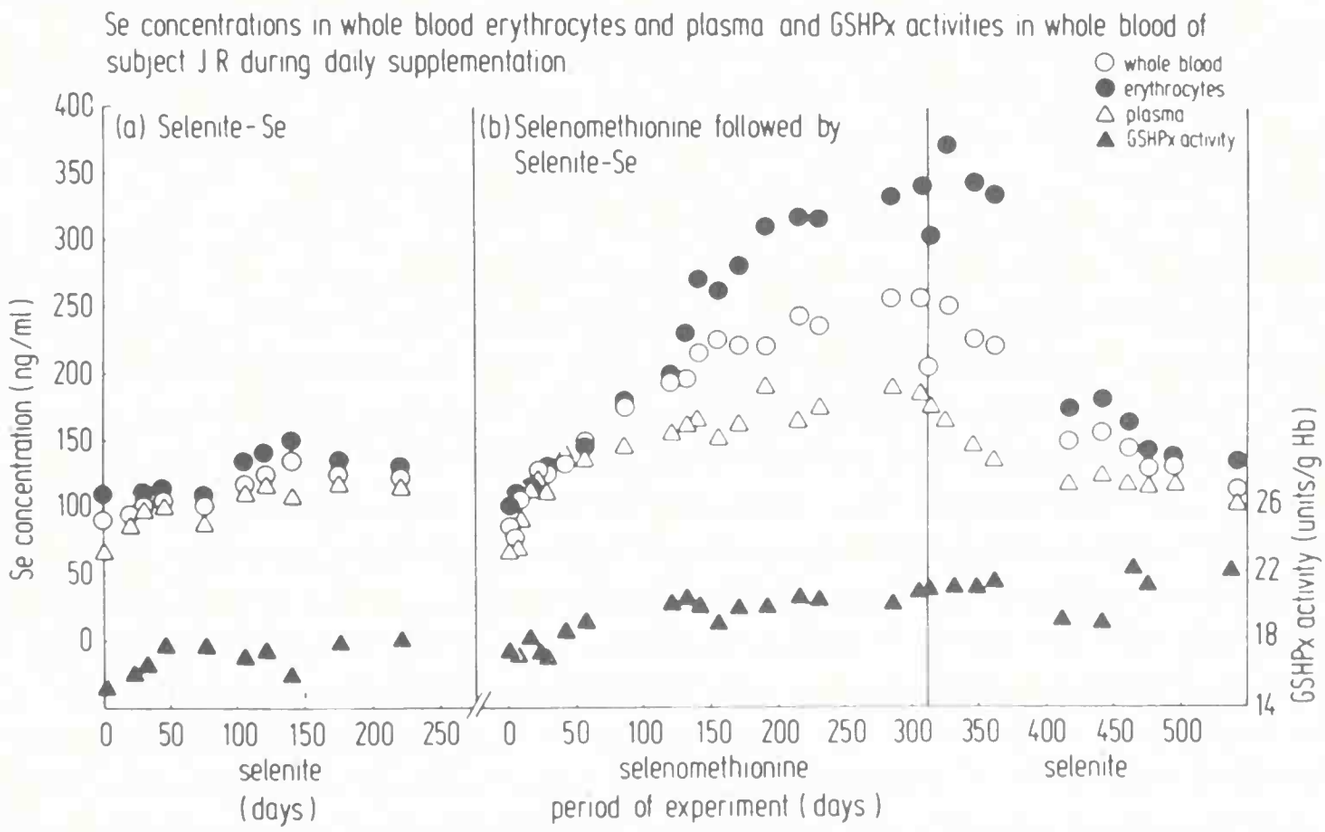


Figure 9–3. Response to daily supplements of 100 μg of selenium during almost 3 years: 31 weeks as selenite; 26 weeks without supplement; 44 weeks selenomethionine; 26 weeks as selenite. (From Robinson and Thomson, 1981; reproduced with permission.)

blood, largely composed of plasma and erythrocytes, has intermediate properties, but suffers from the problem that changes in hemoglobin affect selenium concentration. Good correlations have been observed between dietary selenium intake and 24-hour urinary excretion; however, recent dietary intake greatly affects random urine samples (Robinson et al., 1985).

Repeat Measures

Repeated measures of an indicator over time provide some idea of the within-person variability, and thus the likelihood that the indicator is a stable estimate of long-term intake. If repeated measures of a biochemical indicator vary substantially over time in the same individuals, then a single measure does not reflect true, long-term intake. This lack of consistency may occur because diet has truly changed over the interval between measurements, or because the measure is overly sensitive to short-term influences, such as a recent meal. In either event, the single measurements with that indicator do not provide useful tests of hypotheses related to dietary intake over the period represented by the replicate measurements. If variation between repeated measurements is small, then two possibilities exist. The indicator may be a well-integrated measure of long-term intake, or it may be relatively insensitive to both dietary intake and time. To distinguish between these possibilities, information is required from studies in which intake is known. For example, Taylor and colleagues (1987), in a study of the selenium status of 78 adults evaluated at four 3-month intervals, observed intraclass correlation coefficients (representing the average correlation between repeated measurements) of $r = 0.76$ for serum selenium and $r = 0.95$ for whole blood selenium. This suggests that whole blood selenium may be less susceptible to temporal variation and thus a better index of long-term intake than serum selenium. Additional evidence that whole blood selenium reflects dietary intake was, however, needed before it could be concluded that whole blood selenium is the superior measure.

The importance of adjusting for relevant covariates in calculating correlation coefficients is demonstrated by the data of Gey and colleagues (personal communication) on vitamin E. The unadjusted correlation coefficient between measures of vitamin E separated by 6 years is $r = 0.65$, implying a satisfactory reproducibility. As previously discussed, however, vitamin E concentration is also strongly related to serum cholesterol as it is transported in blood lipoproteins. The correlation over 6 years for repeated measures of cholesterol was also $r = 0.65$. This moderate correlation alone generates some correlation of vitamin E levels over time, even if there was no true correlation of serum vitamin E with dietary intake. In fact, adjusting for serum cholesterol reduced the vitamin E correlation to $r = 0.46$, indicating that a substantial proportion of the correlation for unadjusted vitamin E was due to the fact that serum cholesterol was relatively reproducible over the 6 years. Information on the within-person variance in a measure can be used to correct correlation coefficients or relative risks for attenuation due to random within-person variation in the biochemical parameter (see Chapter 12).

SPECIFIC BIOCHEMICAL INDICATORS FOR SELECTED NUTRIENTS

The remainder of this chapter is devoted to a review of available biochemical indicators for selected nutrients of current epidemiologic interest. In this rapidly evolving field it is obvious that some of the information will be quickly superceded. Before and during any epidemiologic study employing a biochemical measurement, close collaboration with a biochemically oriented scientist performing the assay or closely familiar with the method, therefore, usually provides optimal results as well as stimulating intellectual exchange. The last complete review of this subject appears to have been that of Sauberlich and colleagues (1974). A more recent report of an ad hoc panel assembled by the Federation of American Societies for Experimental Biology (FASEB) to advise the National Center for Health Statistics on the conduct of the Health and Nutrition Examination Surveys (HANES), provides much valuable information (Klasing and Pilch, 1985). A review by Riboli and colleagues (1987) contains much thoughtful detail on methodologic considerations, in addition to specific information on selected indicators of particular interest in cancer research.

The overall approach in this section has been to focus on aspects of particular relevance to epidemiologic studies. These include the nature of the indicator itself, its relation to dietary intake and variation within-person over time, the general technical method of measurement and its error, the stability of the compound in storage, and nondietary factors that may influence levels. For many biochemical indicators it is apparent that this basic information is far from complete, and thus the format and detail varies between indicators. A further reason for variation is that we have attempted to make the depth of coverage proportional to the degree of current epidemiologic interest in each nutrient. Most readers will find this section more useful as a reference than as material to be consumed in its entirety.

VITAMIN A

Vitamin A may be ingested either as a preformed, active vitamin usually as a retinyl ester (which can readily be converted to retinol, the primary form of vitamin A circulating in blood), or as carotenoids (a large family of compounds, some of which have provitamin A activity and can be converted to retinol in the intestinal epithelium). Preformed vitamin A is naturally found only in food from animal sources, whereas carotenoids are primarily obtained from fruits and vegetables, although a small amount is contained in dairy products. Postulated associations between vitamin A consumption and endpoints as various as xerophthalmia, childhood survival (Tarwotjo et al., 1987), and cancer (Sporn and Roberts, 1983), have created considerable interest in the biochemical assessment of vitamin A intake.

Measures

Retinol

Retinol can be accurately measured in serum or plasma; however, these measurements have only a limited interpretability as measures of dietary intake in well-nourished populations. In such populations, the liver stores at least 90 percent of the total body reserve of vitamin A (Olson, 1984) and buffers serum levels over a wide range of dietary vitamin A intake. Thus, serum levels are a poor reflection of vitamin A status unless the liver stores are either very depleted or highly saturated (Russell et al., 1984; Underwood, 1984). When 25,000 IU of retinyl palmitate (13,750 retinol equivalents—a dose more than five times a typical dietary intake) was given daily for 8 weeks to 15 healthy, U.S. adults, no appreciable increase in plasma retinol was observed (Willett et al., 1983a). Wald and colleagues (1985), using supplements of between 10,000 and 36,000 IU of retinyl palmitate daily, demonstrated a small (2%) increase in serum retinol concentrations after 3 months of supplementation. A supplement of 10,000 IU of vitamin A daily for 4 weeks also resulted in a small, statistically significant mean increase in plasma retinol (9% over baseline values) in a group of apparently healthy U.S. women previously identified as having relatively low plasma retinol levels (Willett et al., 1984a). The correlation between preformed vitamin A intake (estimated from a self-administered seimiquantitative food frequency questionnaire) and plasma retinol was weak ($r = 0.17$ adjusted for age, gender, total caloric intake, and plasma lipids) and not statistically significant ($p = 0.12$). Thus, blood retinol measurements may be used in well-nourished populations to identify the rare persons with unusually low intakes, but are likely to misclassify individuals over a wide range of typical intakes. Blood retinol may be a more sensitive indicator of vitamin A status in undernourished populations in which prolonged low intake of vitamin A is common. Patwardhan (1969) demonstrated a strong correlation between mean vitamin A intake and mean serum retinol levels using survey data from 11 developing countries. The mean serum vitamin A levels in Indonesian children with night blindness or Bitot's spots, or both (signs of vitamin A deficiency) were less than two-thirds those of normal control children (Sommer et al., 1980).

Retinyl Esters

Under normal conditions of vitamin A nutriture the concentration of retinyl esters is low (less than 5% of total serum vitamin A) (Smith and Goodman, 1976). When vitamin A intake is excessive, serum retinol may be only slightly elevated, but serum retinyl esters are markedly increased. If hypervitaminosis A is an exposure of interest in an epidemiologic study, then a fasting retinyl ester level is probably the most discriminating test available.

Relative Dose Response Test

The relative dose response (RDR) test is based on the observation that in vitamin A-deficient individuals, retinol-binding protein (RBP) accumulates in the liver (Loerch et al., 1979; Campos et al., 1987). When a small vitamin A challenge is given, RBP-bound vitamin A is rapidly released into the blood. The

method involves taking a blood sample immediately before, and 5 hours after, a small oral dose (e.g., 1500 IU) of vitamin A. The RDR% is calculated as:

$$\text{RDR\%} = \frac{(\text{plasma retinol at 5 hours}) - (\text{plasma retinol at 0 hours})}{(\text{plasma retinol at 5 hours})} \times 100\%$$

Thus, higher values of RDR% are associated with decreased liver retinol stores. In a study of 12 children with liver disease, for instance, Amedee-Manesme and co-workers (1987a) report that five out of five children with RDR% values >20 percent had total liver vitamin A concentrations of less than 20 μg/g, whereas six out of seven children with an RDR% of <10 percent had liver vitamin A values of more than 20 μg/g.

Like serum retinol, however, the RDR test is quite insensitive in well-nourished populations, and its utility in undernourished populations remains to be fully defined (Russell et al., 1984). Conditions that affect protein metabolism (malabsorption, liver disease, severe protein malnutrition) may influence RBP availability, and thus confound the test (Russell et al., 1983).

Serum Retinol Binding Protein

Determinations of serum RBP levels have also been used to assess vitamin A status. Because RBP and vitamin A circulate as a trimolecular complex with transthyretin, little unbound RBP is present in blood. Measurement of RBP on a molar basis should be closely related to blood retinol. Indeed, serum retinol levels are highly correlated to RBP (Willett et al., 1984c; Vahlquist et al., 1978). Because RBP is also sensitive to the disease processes listed previously it may not be a good indicator of vitamin A status when these conditions exist. Due to the dependence of retinol on RBP for transport to tissues, however, this may be a measure of the physiologically available form (Underwood, 1984).

Retinal Rod Function

Tests of night blindness are sensitive measures of vitamin A deficiency (Russell et al., 1984). Classic testing of dark adaptation requires an expensive and delicate dark adaptometer, and is thus unsuitable for large epidemiologic studies. The recently developed rapid dark adaptation test (RDAT) (Thorton, 1977) requires only simple equipment, takes 5 to 15 minutes to perform, and is practical for field studies (Solomons et al., 1982). False-positive results may occur due to congenital night blindness and other nutritional deficiencies that affect dark adaptation (e.g., zinc and protein deficiency). Substantial interobserver variation (Russell et al., 1984) and its inappropriateness for young children limit the use of this functional test in epidemiologic studies.

Impression Cytology

A relatively simple test for detection of severe vitamin A deficiency is the determination that goblet cells are absent from the epithelium of the bulbar conjunctiva (Wittpen et al., 1986; Amedee-Manesme et al., 1987b). In the simplest version of this test, a strip of filter paper is applied to the bulbar conjunctiva, the cells transferred to a glass slide, and the cells fixed, stained, and examined. This test, yet to be extensively evaluated in a large population, offers substantial logis-

tic advantages relative to the collection, storage, and analysis of blood, and may prove to be a useful clinical and research tool in developing countries.

Laboratory Techniques

High-performance liquid chromatography (HPLC), with either ultraviolet or fluorometric detection, is the preferred method for serum retinol determination in microsamples (Bieri et al., 1979; Kalman et al., 1987). Driskell W. and colleagues (1985) note that application of this technique to long-term frozen sera may result in substantial degradation of vitamin A during the procedure, unless a modification of the technique is employed. Although older colorimetric techniques are reasonably comparable to HPLC for the measurement of total vitamin A (Pilch, 1985), HPLC has the advantage that it separates retinol from its esters (Bankson et al., 1986). The between-run CV% for serum retinol determined by HPLC by Biesalski and co-workers (1986) were 3.5 to 3.9 percent. Speek and colleagues (1986) have used an HPLC method to measure retinol in 5-μl aliquots of plasma collected from over 2000 Thai children, and report a within-assay CV% of 3.9 percent and a between-assay CV% of 5.0 percent.

Retinol binding protein may be measured in serum obtained by finger prick using a radial immunodiffusion assay technique (Arroyave et al., 1982), which has the advantage that diffusion plates can be taken into the field for use.

Stability

HPLC recovery of retinol from heparinized plasma is equivalent to that from serum. Oxalate or citrate as anticoagulants, however, reduce retinol concentrations and it is possible that EDTA also causes a slight reduction (Peng et al., 1987; Nierenberg, 1984). Retinol is sensitive to ultraviolet light, thus samples must be protected from any unnecessary sunlight exposure. Retinol and RBP concentrations in samples maintained at 26 to 28°C for 24 hours were 11.9 and 8.4 percent less, respectively, than in samples kept at 4°C (Mejia and Arroyave, 1983). No statistically significant loss of blood retinol occurred in samples stored on ice and kept in the dark for 24 hours before centrifuging (Peng et al., 1987), however, only three samples from each of two subjects were tested. Serum retinol appears to be quite stable when frozen at −80°C for periods up to 7 years (Willett et al., 1984b). No loss of retinol occurred in plasma stored at −20°C for 12 months (Peng et al., 1987). RBP is more stable than retinol during extended storage, and may be a better indicator of initial vitamin A status, especially if storage conditions have been less than optimal (Olsen, 1984).

Other Determinants of Serum Retinol

In addition to a variety of clinical conditions that alter serum retinol levels (Klasing and Pilch, 1985), serum retinol increases with age (Russell et al., 1983). Serum retinol levels also tend to be higher among men, among users of oral contraceptives (Stryker et al., 1988; Bamji and Ahmed, 1978), and, in one report, among smokers (Biesalski et al., 1986). Comstock and colleagues (1988)

observed slightly lower serum retinol levels among male smokers, and slightly higher levels among female smokers. In a study of 201 subjects, plasma retinol was increased among subjects drinking more than 44 g of alcohol per day (Herbeth et al., 1988). Plasma retinol is positively correlated with plasma cholesterol (Willett et al., 1984b; Russell-Briefel et al., 1985), for reasons that are not clear as it is not transported in the lipoprotein particles.

CAROTENOIDS

Although the only established function of carotenoids in humans is their in vivo conversion to vitamin A, Peto and others (1981) have advanced the hypothesis that β-carotene itself may have a protective effect against cancer, independent of any effect mediated through its conversion to vitamin A. Thus, measurement of carotenoids may be of intrinsic interest, as well as providing additional information about vitamin A status.

Carotenoids are red and yellow fat-soluble pigments synthesized in nature exclusively by photosynthetic microorganisms and plants. The basic structure of eight isoprenoid units can be modified by a variety of chemical reactions, producing a group of compounds with a heterogeneous biologic activity. The major carotenoids with vitamin A activity in common food sources are β-carotene, α-carotene, γ-carotene, and cryptoxanthin. Some carotenoid pigments have no detectable vitamin A activity (e.g., lycopene) (Underwood, 1984).

Measures

Plasma and Serum

Blood carotenoid levels are very sensitive to dietary intake as they are not closely regulated by homeostatic mechanisms. A daily β-carotene supplement of 30 mg (5000 retinol equivalents) approximately tripled plasma total carotenoid levels after 8 weeks (Willett et al., 1983a). Among individuals given daily β-carotene supplements, blood levels rose progressively over 2 to 3 weeks and declined over a similar period when supplementation stopped (M. Stampfer, personal communication). Among nonsupplement users, blood levels vary remarkably among both men and women; for both groups the between-person coefficient of variation is more than 100 percent (Stryker et al., 1988). Among 24 male subjects who provided four consecutive weekly blood specimens, the within-to-between-person variance ratio for plasma β-carotene level was 0.62 (intraclass r = 0.79) (Tangney et al., 1987). This sensitivity to intake and the capacity to integrate over a period of several weeks makes even a single measurement of blood β-carotene a potentially good index of dietary intake.

Adipose Carotenoid Concentration

Among 19 adults who had 1 to 2 g of abdominal adipose tissue removed at surgery, total carotenoid concentration varied 40-fold between individuals (Parker, 1988). Beta-carotene was 20.2 percent and lycopene was 18.5 percent, of total carotenoids. The high between-person variability in carotenoid concentration

suggests that adipose tissue measurement may be useful, but correlations with carotenoid intake have not been reported.

Laboratory Techniques

Total carotenoids may be measured spectrophotometrically and by HPLC (Bieri et al., 1985). The HPLC method differentiates α-carotene, β-carotene, cryptoxanthin, and lycopene from other carotenoids, the sum of these peaks accounting for more than 90 percent of the total carotenoids in a total lipid extract. The between-assay CV% for β-carotene measured by HPLC on consecutive days by Vuilleumier and colleagues (1983a) was 3 percent.

Storage

β-carotene, like retinol, is sensitive to light and heat. β-carotene levels declined by 6.7 percent per day in blood samples kept in the dark at room temperature for 2 days (S. Hankinson, personal communication). Samples stored at $-20°C$ deteriorated by approximately 15 percent over 6 months; thus storage at $-70°C$ is necessary. Qualitatively distinguishable peaks for α-carotene and β-carotene were present in serum frozen at $-70°C$ without thawing for 9 years (Willett et al., 1984c) and the levels of total carotenoids were not markedly lower than expected (Willett et al., 1984b), however, the quantitative stability of carotenoids during long-term storage remains to be determined.

Other Determinants of Carotenoids

The mean level of β-carotene was approximately 20 percent higher among women in Washington County, Maryland than the level observed among men, and there was a small increase with age (Comstock et al., 1988). Smoking and alcohol consumption are independently associated with reduced serum β-carotene levels in both sexes (Stryker et al., 1988) and these effects do not appear to be explained entirely by differences in dietary carotene intake. In multiple regression analyses controlling for dietary carotene, smoking, and alcohol consumption, these investigators observed a significantly positive association of β-carotene with plasma cholesterol, and a significantly inverse relationship with plasma triglycerides. In the Basel cohort, plasma β-carotene levels tended to be higher from August to November, and lower from January to May, presumably reflecting seasonal changes in diet (Gey et al., personal communication). More extreme seasonal variations may be seen in developing countries (see Chapter 3).

VITAMIN E

Vitamin E is a powerful intracellular antioxidant that reduces lipid peroxidation, and lowers transformation frequencies in in vitro carcinogenesis assays. In humans, clinical vitamin E deficiency has been reported only in premature

infants, newborns, and persons with fat malabsorption disorders. The two principal forms of vitamin E are α-tocopherol and γ-tocopherol, of which the former has the greater biologic activity. In a cross-sectional study, serum α-tocopherol concentrations were between 5.6 and 8.3 times higher than γ-tocopherol (Behrens and Madere, 1986). α-tocopherol is transported in blood as part of a lipoprotein complex, mainly in association with low density lipoprotein (LDL)–cholesterol. Unlike vitamin A, there is no principal storage organ for vitamin E.

Measures

Plasma and Serum

The most common tests of vitamin E nutriture have been measurement of either α-tocopherol or total vitamin E in blood. The major problem with interpretation of these results is confounding by blood lipid concentration, due to the strong positive correlation between vitamin E levels and serum cholesterol and total lipid concentrations. Willett and colleagues (1983b) calculated correlations between plasma α-tocopherol and total cholesterol ($r = 0.59$, $p = \leq 0.001$), high density lipoprotein (HDL) ($r = -0.02$), and LDL ($r = 0.56$, $p \leq 0.001$). A concurrent measurement of either serum cholesterol or LDL permits adjustment of the α-tocopherol measurement; without this, considerable misclassification results.

Plasma levels are moderately responsive to α-tocopherol intake. A daily α-tocopherol supplement of 800 IU (approximately 100 times usual dietary intake) approximately doubled plasma α-tocopherol levels in well-nourished adults after 16 weeks (Willett et al., 1983a). The simple correlation between vitamin E intake estimated from a self-administered semiquantitative food frequency questionnaire and plasma α-tocopherol was $r = 0.12$ ($p = 0.19$). Adjustment for total plasma cholesterol increased the correlation to $r = 0.28$ ($p = 0.02$), and adjustment for age, sex, total caloric intake, and plasma triglycerides further increased the partial correlation to 0.34 ($p = 0.006$). Knekt and colleagues (1988) observed a correlation of $r = 0.65$ between measures of vitamin E obtained 4 years apart from 105 adults in Finland. Over a 6-year period, Gey and co-workers (personal communication) also found a correlation of $r = 0.65$ for repeated measures. Lipid-adjustment, however, reduced this correlation to $r = 0.46$, implying that part of the original correlation was determined by the fact that cholesterol measurements were also highly correlated over the 6-year period, which exaggerated the correlation for total vitamin E. Nevertheless, it does appear that a single measurement of plasma α-tocopherol can represent long-term vitamin E intake to a reasonable degree, reflecting both a capacity to integrate intake over a few weeks and an element of long-term stability of diet among persons.

Erythrocyte Tocopherol

α-Tocopherol and its β, δ, and γ analogues can be measured in red blood cells by HPLC. The α-tocopherol content of both red blood cells and platelets increased in a dose-response manner among 20 subjects consuming 41 IU per day, and then 136 IU per day, for two 6-week periods (Lehmann et al., 1988).

Among 261 healthy Japanese children, red cell α-tocopherol was highly correlated with plasma α-tocopherol ($r = 0.59$, $p < 0.001$). It has been suggested that red blood cell α-tocopherol may be a more appropriate measure of bioavailable vitamin E than plasma α-tocopherol, particularly if lipid-adjustment is not possible (Mino and Nagamato, 1986). The exact nature of the relationships between vitamin E intake, plasma and red blood cell tocopherol, and plasma lipids, however, remain to be determined.

Adipose Tissue Tocopherol

As much as 90 percent of body tocopherol is contained in adipose tissue, and large oral doses of vitamin E increase tocopherol levels in adipose tissue obtained from patients with abetaliproteinemia, whose initial levels are low (Traber and Kayden, 1987). Few data are available on the relation between adipose tocopherol and vitamin E intake over the normal range. A recent report of substantial variability in α-tocopherol/cholesterol ratios between needle fat aspirates from different anatomic sites suggests that standardization of needle aspirate site may be advisable (Handleman et al., 1988)

Erythrocyte Hemolysis Test

The ability of erythrocytes to resist hemolysis when stressed with hydrogen peroxide is a test of the antioxidant protection provided by vitamin E to the erythrocyte membrane (Russell et al., 1984). The rate of hemolysis is inversely correlated with serum tocopherol levels; however, deficiencies of other nutrients that play a role in cellular antioxidant function, such as selenium, may also influence the hemolysis rate (Klasing and Pilch, 1985). The requirement for fresh erythrocyte samples also limits the applicability of this test in many epidemiologic studies.

Pentane–Ethane Breath Test

The concentrations of pentane and ethane (products of the peroxidation of linolenic and linoleic acids) may be measured in exhaled breath by gas chromatography, and high levels may indicate vitamin E deficiency (Russell et al., 1984). Among 10 adult subjects, pentane output was negatively correlated with plasma vitamin E levels ($r = -0.66$, $p < 0.01$), and breath pentane was significantly reduced in 5 subjects supplemented with 1000 IU of α-tocopherol for 10 days (Lemoyne et al., 1987). Application of this test in large epidemiologic studies is obviously problematic.

Laboratory Techniques

Although vitamin E may be measured by colorimetric or spectrofluorometric techniques, HPLC is probably the method of choice. HPLC allows the differentiation of α-tocopherol and γ-tocopherol (Behrens and Madere, 1986) and the simultaneous determination of serum retinol (Bieri et al., 1979). Biesalski and colleagues (1986) report CV% between assays of 4 to 4.5 percent.

Storage and Stability

In a study of 12 adult subjects, α-tocopherol decreased by 1.6 percent per day, and γ-tocopherol increased by 4.9 percent per day, in blood stored at room temperature in the dark for 2 days (S. Hankinson, personal communication). Biesalski and colleagues (1986) comment that α-tocopherol is stable for more than 2 weeks when stored at -34°C in the dark. Willett and colleagues (1984b) report that after more than 7 years of storage at -70°C, the mean α-tocopherol among 321 participants in the HDFP was 1.26 mg/dl compared with 1.02 mg/dl in samples of freshly drawn blood from Boston residents, providing indirect evidence that loss during storage was minimal.

Measurement of Other Determinants

When assessing serum or plasma vitamin E levels, it is desirable to obtain concurrent measurements of serum cholesterol or total lipids. In multiple regression analyses controlling for plasma cholesterol, neither age nor sex were significant predictors of plasma α-tocopherol (Willett et al., 1983b). Cigarette use and alcohol consumption also do not appear to be important determinants of plasma α-tocopherol levels (Stryker et al., 1988; Comstock et al., 1988).

VITAMIN D

Epidemiologic studies of vitamin D status in children have traditionally been concerned with rickets, now rare in developed countries subsequent to the fortification of milk. Recent interest in vitamin D nutrition in adults centers on bone mineralization, osteoporosis, and osteomalacia. Postulated associations between calcium and hypertension (Sowers et al., 1985), and calcium and colorectal cancer (Garland et al., 1985), have led to investigations of the role of vitamin D in the etiology of these diseases.

Vitamin D may be absorbed from the diet or synthesized in the sun-exposed skin. The relative contributions of these two sources varies. Clinical vitamin D deficiency is unlikely, however, unless the supply from both sources is reduced (Parfitt et al., 1982). Vitamin D is synthesized in the skin by the ultraviolet-catalyzed conversion of 7-dehydrocholesterol to vitamin D_3 (cholecalciferol), which is slowly released into the bloodstream bound to an α-globulin (transcalciferin). Natural sources of dietary vitamin D are few, however, many countries fortify margarine or dairy products either with vitamin D_2 (ergocalciferol) or vitamin D_3. Plasma derivatives of vitamin D_2 are thus of exogenous origin only, whereas derivatives of vitamin D_3 may arise from either diet or skin (Parfitt et al., 1982). The metabolic fate of both these compounds is the same.

Vitamin D must be metabolically activated before functioning biochemically, initially by 25-hydroxylation in the liver to form 25-hydroxyvitamin D [25(OH)D], the major circulating metabolite, and then by 1-α-hydroxylation in the kidney to produce 1,25-dihydroxyvitamin D [1,25$(OH)_2$D], which is the

active metabolite. This last step is the principal site of regulation of vitamin D metabolism. Formation of $1,25(OH)_2D$ is increased by low serum phosphate levels, and by low serum calcium mediated by parathyroid hormone. Correction of these deficits reduces the stimulus for 1-α-hydroxylation, thus closing the feedback loop (Haussler and McCain, 1977; DeLuca, 1979). Cellular receptors for $1,25(OH)_2D$ have been found not only in the intestine and bone, but in many other tissues, suggesting that $1,25(OH)_2D$ is fundamental to the metabolism of many cell types, probably because of its role in intracellular calcium regulation (Dickson, 1987).

Measures

25-Hydroxyvitamin D [25(OH)D]

The most useful measure of vitamin D status is the plasma level of the major circulating metabolite, 25-hydroxyvitamin D (Parfitt et al., 1982; Russell et al., 1984). In six normal subjects, more than 80 percent of the total 25(OH)D existed in the form $25(OH)D_3$ (Haddad and Hahn, 1973). Substantial seasonal variation exists, especially in elderly populations, with plasma 25(OH)D levels reaching a nadir in winter and rising during the summer, presumably reflecting seasonal ultraviolet-sunlight exposure. Values are higher in men than women, and higher in vitamin D supplement users (Omdahl et al., 1982). In an international correlation study serum 25(OH)D was inversely related to latitude, and positively correlated with mean vitamin D intake (McKenna et al., 1985). In countries with vitamin D fortification of foods, the population most likely to be both deprived of sunlight and inadequate consumers of dietary vitamin D, is the elderly. Significantly lower plasma 25(OH)D concentrations in elderly subjects compared with younger controls have been demonstrated in several studies (Omdahl et al., 1983; Parfitt et al., 1982).

In a study of 373 healthy, noninstitutionalized women, Sowers and colleagues (1986) observed a weak ($r = 0.11$) positive correlation between estimated vitamin D intake from food (based on a 24-hour recall), and serum 25(OH)D. This correlation improved to $r = 0.24$ when vitamin D from supplements was included in the diet score. The correlation between estimated sunlight exposure and serum 25(OH)D was $r = 0.26$. Newton and colleagues (1985) investigated the relation between vitamin D intake and plasma 25(OH)D among 57 institutionalized elderly women (whose sunlight exposure was minimal). Both vitamin D_2 and vitamin D_3 intake were strongly correlated with plasma $25(OH)D_2$ and $25(OH)D_3$, respectively, as were total vitamin D intake and total plasma 25(OH)D ($r = 0.55$). In a study of 125 patients with hip fractures (Lips et al., 1987), a significant correlation was observed between vitamin D intake and serum 25(OH)D only among patients with low sunshine exposure ($r = 0.54$). These data suggest that sunshine exposure is the most important determinant of total 25(OH)D levels, and that total 25(OH)D may only be a good marker of dietary intake in subjects with low sun exposure.

In patients with hypervitaminosis D, plasma 25(OH)D levels were approximately 15 times higher than those of normal controls (Hughes et al., 1976). As this was exogenous vitamin D, most of the plasma metabolites were of the

$25(OH)D_2$ form. Hypercalcemia frequently coexists with high levels of 25(OH)D and is a useful marker of hypervitaminosis D (Klasing and Pilch, 1985).

1,25-dihydroxyvitamin D [$1,25(OH)_2D$]

Although measurement of $1,25(OH)_2D$ is a direct measurment of vitamin D hormone activity, it is not a good reflection of vitamin D nutritional status because of the influences of calcium and phosphate levels, and parathyroid hormone, described previously.

Serum Alkaline Phosphatase Activity

Alkaline phosphatase activity was used as an indirect measure of vitamin D status before the availability of direct measurements of vitamin D metabolites. Alkaline phosphatase activity increases in proportion to vitamin D deficiency, however, it is susceptible to other disease processes such as Pagets' disease, hyperparathyroidism (increased), and protein-energy malnutrition (decreased), which may confound the relation with vitamin D (Sauberlich et al., 1974).

Methods

Until recently, a competitive protein-binding assay was widely used to measure 25(OH)D (Belsey et al., 1974). Measurement of $1,25(OH)_2D$, and distinction between $25(OH)D_2$ and $25(OH)D_3$, can be achieved chromatographically (Omdahl, 1978). HPLC for measurement of vitamin D metabolites is now the method of choice (Klasing and Pilch, 1985).

Storage

25(OH)D concentrations were significantly reduced by 9.3 percent in plasma stored for 11 months at $-18°C$, suggesting that long-term storage may not be feasible if vitamin D analyses are desired (Norris et al., 1986), or that colder temperatures are needed.

VITAMIN K

Vitamin K is essential for the activation of four proteins in the clotting cascade—prothrombin, factor VII, factor IX, and factor X. Deficiency results in a hemorrhagic tendency. About half of the human vitamin K requirement is derived from green leafy vegetables, meat, and dairy products, and the remainder from biosynthesis by the intestinal flora (Olson, 1980). As vitamin K is widely distributed in food sources, and endogenously synthesized in the gut, vitamin K deficiency is unusual in healthy adults and vitamin K status is a poor indicator of vitamin K intake. Newborn infants, born without intestinal flora, and adults with fat malabsorption syndromes, liver disease, inflammatory bowel disease, or taking gut-sterilizing antibiotics, may be prone to vitamin K deficiency.

Vitamin K status has traditionally been determined by functional methods

that measure prothrombin activity. Recent development of immunochemical assays has made direct measurement of prothrombin in plasma feasible.

Functional Assays

Of the many tests for bleeding abnormalities, the most specific for vitamin K deficiency is the prothrombin time, which tests the time plasma takes to clot when excess calcium and thromboplastin are added. The prothrombin time does not become prolonged until circulating prothrombin is less than 40 percent of normal. Thus, it is not a good test for mild vitamin K deficiency (Russell et al., 1984).

Direct Assays

Recently HPLC has been used to measure vitamin K in human plasma (Shearer et al., 1982), although this required 7 to 10 ml of plasma. Immunoassays for the direct measurement of both normal prothrombin, and the abnormal forms that appear during vitamin K deficiency (Blanchard et al., 1981), are much more sensitive indicators of mild vitamin K deficiency than the prothrombin time. These assays can be performed on submilliliter amounts of plasma. They are, however, not yet commercially available and require a considerable investment of time and laboratory expertise.

Measurement of Other Determinants

Information about disorders interfering with endogenous synthesis (e.g., bowel disease) or requiring chronic antibiotic treatment, is useful in the interpretation of vitamin K status (Olson, 1987).

VITAMIN C

Vitamin C occupies a special place in the history of nutritional epidemiology because the discovery that scurvy was preventable by dietary manipulation remains a classic demonstration of the relation between a specific dietary deficiency and a specific disease. Indeed, James Lind's study (Lind, 1753) of the treatment of scurvy was one of the first intervention studies performed (MacMahon and Pugh, 1970), although the sample size he used (two patients for each treatment) was small by modern standards! Recent enthusiasm for high-dose administration of vitamin C as a treatment for cancer (Cameron and Pauling, 1978), and the suggestion that low vitamin C intake may predispose to cancer (Cameron et al., 1979), has led to renewed epidemiologic interest in this vitamin. Although experimental data suggest many mechanisms for a protective effect of vitamin C for cancer, epidemiologic evidence is sparse (Committee on Diet, Nutrition and Cancer, 1982; Willett and MacMahon, 1984).

The pharmacokinetics of vitamin C are well described. Over the range of

usual dietary intake, vitamin C is readily absorbed from the small intestine, the absorption efficiency being 90 percent or more (Olson and Hodges, 1987). The unadjusted correlation coefficients between questionnaire-derived dietary ascorbic acid intake and plasma and leukocyte ascorbic acid concentrations in a heterogenous population were $r = 0.43$ and $r = 0.31$, respectively (Loh, 1972). Vitamin C is not protein-bound, and thus circulates freely between plasma and tissues, with the tissue concentrations being 3 to 10 times higher (Olson and Hodges, 1987). As ascorbic acid intake increases, plasma, serum, buffy coat (platelets and leukocytes), and leukocyte levels increase monotonically, although the rate of increase levels off as intake approaches the upper end of the normal range (Basu and Schorah, 1982; Jacob et al., 1987). When dietary vitamin C is completely eliminated, ascorbic acid becomes undetectable in plasma after 35 to 40 days, in whole blood after 80 to 90 days, and in leukocytes after 100 to 120 days (Baker et al., 1980). Thus, leukocyte vitamin C is probably a preferable measure of long-term intake, whereas plasma and serum levels reflect more recent intake (Jacob et al., 1987). This sensitivity to short-term intake was reflected in the Basel study in which the correlation between plasma vitamin C values in men obtained 6 years apart ($r = 0.28$) was much lower than the correlations obtained for vitamin E and β-carotene (Gey, personal communication). This suggests that a single plasma vitamin C determination will misclassify many individuals with respect to their actual long-term intake. Despite this, the much greater expertise and expense needed to separate and store leukocytes, not to mention the larger initial blood sample size required, means that plasma or whole blood are probably more appropriate for large epidemiologic studies. The best method for measuring vitamin C nutritional status is total body pool estimation using isotope dilution or excretion techniques. Although inapplicable to large studies, they remain the standard for evaluating other methods.

Urinary ascorbate is a less sensitive measure than plasma over the normal range of vitamin C intake, but as renal clearance increases at high intakes (when the plasma values plateau), urinary measures may identify megadose vitamin C supplement users. It has been suggested that 24-hour urinary levels of an ascorbic acid metabolite, ascorbitol, may reflect body pool size (Klasing and Pilch, 1985).

Methods

The classic method for blood and serum vitamin C measurement has been derivatization with 2,4-dinitrophenylhydrazine with colorimetric analysis. HPLC methods are also becoming available, but remain to be fully proven in biologic specimens (Pachla et al., 1985).

Stability

Without special preservation, vitamin C deteriorates rapidly during frozen storage (Basu and Schorah, 1982). Plasma samples should be acid-stabilized and frozen at $-70°C$.

Measurement of Other Determinants

Ascorbic acid levels are higher in the plasma and leukocytes of vitamin C supplement users. Leukocyte ascorbic acid is reduced in oral contraceptive users (McLeroy and Schendel, 1973). Several studies have demonstrated lower plasma and leukocyte vitamin C levels among cigarette smokers (Brook and Grimshaw, 1968). Gey and collegues (personal communication) identified seasonal variations in plasma vitamin C similar to, but weaker than, those observed for β-carotene. Acute and chronic infections also lower plasma vitamin C levels (Irwin and Hutchins, 1967).

VITAMIN B_1 (THIAMIN)

Vitamin B_1 deficiency is the cause of beriberi, which principally occurs in regions where polished rice is the staple food. In developed countries, improved nutrition and thiamin fortification of flour has made clinical vitamin B_1 deficiency rare, except among patients with chronic gastrointestinal syndromes and chronic alcoholics, in whom thiamin deficiency may cause Wernicke's encephalopathy. Iber and colleagues (1982) have estimated that 5 percent of U.S. adults over 60 years old have impaired thiamin status, and that the prevalence is higher among the poor, the chronically ill, and the institutionalized. The principal dietary thiamin sources are cereals, legumes, other fresh vegetables, pork, and beef.

The best method for assessing thiamin nutriture in the general population is the measurement of the stimulation of erythrocyte transketolase by thiamine pyrophosphate (TPP) (Sauberlich et al., 1974). Transketolase is an enzyme in the pentose phosphate pathway, and requires thiamin pyrophosphate. The assay measures transketolase activity before and after the addition of TPP, and a large enhancement of enzyme activity by TPP implies a relative deficiency of endogenous thiamin (Sauberlich et al., 1974). Recent improvements in the precision of this test, and the development of automated procedures, make this method more applicable to large studies (Sauberlich, 1984). Basu and colleagues (1974) have developed a colorimetric method to measure transketolase activity in as little as 50 μl of whole blood.

Although HPLC is a precise technique for measuring vitamin B_1 in plasma and whole blood (Bettendorff et al., 1986), this is not a useful measure of thiamin status because even in deficiency the reduction in blood thiamin is small. Similarly, urinary thiamin is not a good index of thiamine nutriture over the normal range of intake, although it may be useful to confirm clinically suspected thiamin deficiency (Sauberlich, 1984).

Methods

Automated methods of measuring erythrocyte transketolase stimulation provide reproducible and sensitive results that correlate well with manual methods (Waring et al., 1982). The coefficient of variation between repeated determinations using a manual method was 5.3 percent (Vuilleumier et al., 1983b).

Storage and Stability

Long-term stability of erythrocyte transketolase is uncertain.

Measurement of Other Determinants

Subjects with diabetes mellitus and polyneuritis have low erythrocyte transketolase activity independent of thiamin intake, whereas activity is increased in patients with pernicious anemia (Kjosen and Seim, 1977). Heavy alcohol intake not only inhibits thiamin absorption but affects its metabolism (Iber et al., 1982).

VITAMIN B_2 (RIBOFLAVIN)

Clinical riboflavin deficiency is rare in developed countries; the principal dietary sources are dairy foods, meat, and riboflavin-enriched flour.

The best indicator of riboflavin nutritional status is the erythrocyte glutathione reductase (EGR) assay. This measures the increase in activity of the enzyme glutathione reductase after the in vitro addition of flavin adenine dinucleotide (FAD). This is expressed as the erythrocyte glutathione reductase assay coefficient (EGRAC), which is the ratio of enzyme activity with and without added FAD (Tillotson and Baker, 1972). In normal subjects little increase occurs, whereas in those with an inadequate intake of riboflavin, the increase is marked, and guidelines for the interpretation of EGRAC have been given (McCormick, 1985).

In a study of the behavioral effects of riboflavin restriction (Sterner and Price, 1973), the EGRAC rose monotonically with increasing length of riboflavin depletion, then returned to baseline during the repletion period. EGRAC was a more stable index of riboflavin status than urinary riboflavin, which displayed more within-person variation. Confirming a high EGRAC value with a low urinary riboflavin value has been recommended for the accurate diagnosis of riboflavin deficiency (Horwitt, 1987). Newer methods for direct measurement of riboflavin in blood have yet to be extensively validated (Sauberlich, 1984).

Methods

For the EGR assay erythrocytes are separated soon after blood collection, and then may be frozen. Only a small sample is required and automated colorimetric methods are available (Sauberlich, 1984).

Measurement of Other Determinants

The EGR assay is not a good measure of riboflavin nutriture in persons with glucose-6-phosphate dehydrogenase deficiency (Thurnham, 1972) and certain endocrine disorders.

VITAMIN B_6

The term *vitamin* B_6 embraces a family of compounds with a structural relationship to pyridoxal phosphate. This enzyme has a function in over 60 different enzyme systems, most of which are involved in amino acid and protein metabolism (Sauberlich et al., 1974). Frank clinical vitamin B_6 deficiency is uncommon, although among at-risk groups such as the elderly (Driskell J., 1978) and adolescent girls (Driskell J. et al., 1985), inadequate vitamin B_6 status, measured biochemically, may be more prevalent. Low vitamin B_6 intake has been implicated in the etiology of cancer (Thanassi et al., 1985) and coronary heart disease (Willett, 1985).

Measures

Several different approaches to vitamin B_6 status assessment exist, and the most physiologically appropriate measure has yet to be determined (Reynolds and Leklem, 1985). The most common procedure is the direct measurement of plasma pyridoxal 5′-phosphate (PLP). PLP is the major circulating and coenzymatic form of vitamin B_6, comprising over half the total body pool (Lumeng et al., 1985). In an experiment in which subjects ingested constant diets that were normal, restricted, or supplemented with vitamin B_6, plasma PLP discriminated well between the different dietary regimens, and there was low within-person fluctuation of plasma PLP. A dose-response relation exists between increasing vitamin B_6 supplement dose and plasma PLP (Schultz and Leklem, 1985). Response to oral vitamin B_6 supplementation in prevously unsupplemented subjects, however, is rapid (Thakker et al., 1987), suggesting that PLP is a better measure of short-term rather than long-term intake. This may be the reason that only modest correlations were found by Willett (1985), between a food-frequency questionnaire estimate of vitamin B_6 intake over 1 year, and a single fasting plasma PLP level among 280 healthy men and women ($r = 0.37$ for men and $r = 0.39$ for women).

Vitamin B_6 measurements in random morning urine samples have been found to correlate reasonably well with 24-hour excretion (Sauberlich et al., 1974). Urinary vitamin B_6 may be even more responsive to recent dietary intake than plasma PLP, however, and thus a poorer indicator of long-term intake. The measurement of the principal metabolic end-product of vitamin B_6, 4-pyridoxic acid (4PA), has also been found to reflect vitamin B_6 intake, however, this requires a 24-hour urine collection (Schultz and Leklem, 1981).

The measurement of erythrocyte transaminase enzymes, such as glutamate-oxaloacetate transaminase (GOT), provide a limited indication of vitamin B_6 status (Sauberlich et al., 1974). A stimulation test measuring the in vitro effect of adding pyridoxal 5′-phosphate also lacks sensitivity and specificity, although recent improvements in methods and the development of automated procedures (Sauberlich, 1984), suggest this assay may hold promise. Another functional test, the tryptophan load test, measures the excretion of tryptophan metabolites in the urine after a tryptophan load, testing the integrity of the tryptophan meta-

bolic pathway, which requires PLP as a coenzyme for several reactions. Many factors other than vitamin B_6 intake affect tryptophan metabolism, and the necessity for a 24-hour urine collection makes this test unwieldy for large field studies.

Methods

Several well-validated chemical and enzymatic assays for measuring PLP in plasma exist (Sauberlich, 1984). HPLC procedures have the advantage that they can be used to separate and quantitate the different vitamin B_6 compounds. Vitamin B_6 measurements in finger prick blood samples correlated well with measurements using venous blood (Andon and Reynolds, 1987).

Storage

Howard and colleagues (1984) found PLP levels to be highly stable in plasma frozen at −20°C over a 700-day period. Based on 113 determinations of aliquots from a stored pool over this interval, the rate of decline was only 2.2 percent per year. Camp and co-workers (1983) found PLP to be stable in plasma stored at −80°C for at least 10 days.

Measurement of Other Determinants

Willett (1985) observed little relation between plasma PLP and age, gender, or obesity. A transient increase in plasma PLP and urinary vitamin B_6 occurs after strenuous physical activity (Leklem, 1985; Manore et al., 1987) and women using oral contraceptives or postmenopausal estrogens may be at higher risk of developing vitamin B_6 deficiency (Miller, 1985). Certain drugs and ethanol also lower plasma PLP levels (Bhagavan, 1985).

FOLACIN

Folacin is a generic term for compounds with chemical structures and nutritional functions similar to folic acid (also called pteroylglutamic acid). Folacin analogues function as coenzymes in the metabolism of single-carbon compounds, in particular, nucleic acid synthesis and amino acid metabolism. The principal dietary sources of folacin are leafy vegetables, fruits, cereals, and tea. Between one-half and two-thirds of ingested folacin is absorbed (Herbert, 1987).

The principal manifestation of folate deficiency is megaloblastic anemia. Folate deficiency in the periconceptual period has been suggested as a cause of neural tube defects in infants (Smithells et al., 1983). In the second National Health and Nutritional Examination survey, 1976–1980 (NHANES II), the greatest prevalence of low folate status was among women of child-bearing age (low red blood cell folate in 13%), whereas the prevalence of low levels in children 6 months to 9 years was only 2 percent. The prevalence in older adults (45 to 74 years) was intermediate (8% for men, 4% for women) (Senti and Pilch,

1984). Folate deficiency and megaloblastic anemia are more common in developing countries (Sauberlich et al., 1974).

Measures

The most commonly employed assays of folacin nutriture are serum and red blood cell (RBC) folate levels. Serum levels of folate decline 3 weeks after the initiation of a low folate diet, however, RBC folate remains in the normal range for approximately 17 weeks (Herbert, 1987). The decrease in RBC folate coincides closely with the depletion of liver folate stores (liver is the major storage organ for folate), and the onset of morphologic abnormalities in erythrocytes. This relative insensitivity to short-term changes in intake, and the close relationship with the hematologic manifestations of folate deficiency, suggest that RBC folate is the superior measure of folacin status for epidemiologic purposes. RBC folate correlates well with dietary folate intake (Bates et al., 1982) ($r = 0.51$, $p = 0.02$, among 19 elderly subjects).

An alternative approach to folacin status assessment is the measurement of polymorphonuclear (PMN) leukocyte lobe counts. Hypersegmented neutrophils appear in the bone marrow after about 5 weeks of folate deficiency, and in the blood after about 7 weeks (Herbert, 1987). Hypersegmentation is commonly assessed as the percentage of PMN leukocytes in peripheral blood smears with five or more lobes, or for a more rapid assessment, the presence or absence of a single six-lobed cell. PMN leukocyte hypersegmentation may also be due to vitamin B_{12} deficiency, and is an unreliable indicator of folacin deficiency during pregnancy (Sauberlich, 1984).

Methods

The two principal approaches to folacin measurement in serum and RBCs are microbiologic and radiodilution assays. The microbiologic assay, using the growth dependency of *Lactobacillus casei* on folacin, is regarded as the standard procedure for folacin measurement in biologic specimens. A variety of radioassays are available in commercial kit form, although there is some concern over their validity (Sauberlich, 1984). Both methods were used in NHANES II, and both were subject to relatively high within-run CV%s (of the order of 10% to 20%) (Senti and Pilch, 1984). Independent quality control and validation are, therefore, desirable for investigators using either method.

Storage

The effect of storage on measures of folacin status is uncertain.

Measurement of Other Determinants

Serum and RBC folate are reduced in smokers and in many pregnant or highly parous women ($\geq$ three live births), and are increased in multivitamin supple-

ment users (Senti and Pilch, 1984). Alcohol and prescription drugs (e.g., folate antagonists, phenytoin, and oral contraceptives) may antagonize folacin absorption or utilization. RBC folate may also be reduced in patients with a primary vitamin B_{12} deficiency (pernicious anemia) (Sauberlich et al., 1974).

SELENIUM

Selenium has attracted much scientific attention, due principally to its possible role as a cancer preventive agent (Schrauzer et al., 1977; Clark, 1985; Willett, 1986). Low selenium intake has been linked with a regional cardiomyopathy in China (Keshan Disease Research Group of the Chinese Academy of Medical Sciences, 1979).

Measures

As previously described, several lines of evidence indicate that selenium may be meaningfully measured in blood components and that erythrocyte selenium is probably superior to serum selenium as a measure of long-term intake. The antioxidant function of the selenium-dependent enzyme glutathione peroxidase, measured in plasma, serum, erythrocytes, platelets, or whole blood, may be a useful functional test of selenium status. Glutathione peroxidase activity is reduced among subjects with low serum selenium (Thomson et al., 1977), and increases with selenium supplementation if a deficiency exists (Thomson et al., 1985). Enzyme activity, however, plateaus in selenium replete individuals, and hence is a poor measure of selenium intake among persons with moderate and high exposure. Thus, a linear relation exists between erythrocyte selenium and glutathione peroxidase activity among New Zealand residents, but the two are essentially uncorrelated among short-term visitors to the country (Rea et al., 1979) (Fig. 9–4).

Hair selenium is well correlated with blood selenium in China (Chen et al., 1980). Selenium-containing shampoos, however, represent a major source of environmental contamination in developed countries. Toenail clippings are less exposed to environmental contamination (at least in countries where wearing shoes is the norm), and present major logistic advantages in ease of specimen collection, transport, and storage, compared with blood or urine. Toenail selenium levels appear to reflect selenium intake. In a study of South Dakotans with high dietary selenium intake, a good correlation was observed between serum and toenail selenium levels ($r = 0.66$) (Longnecker et al., 1987). As toenails grow slowly and are different lengths, if toenails are clipped from all toes simultaneously, these specimens provide a time-integrated measure of intake over much of the preceding year. The feasibility of collecting toenails from large numbers of subjects has been demonstrated in the Nurses' Health Study, in which two-thirds (more than 68,000 women) of the cohort responded to a mailed request to return a set of toenail clippings.

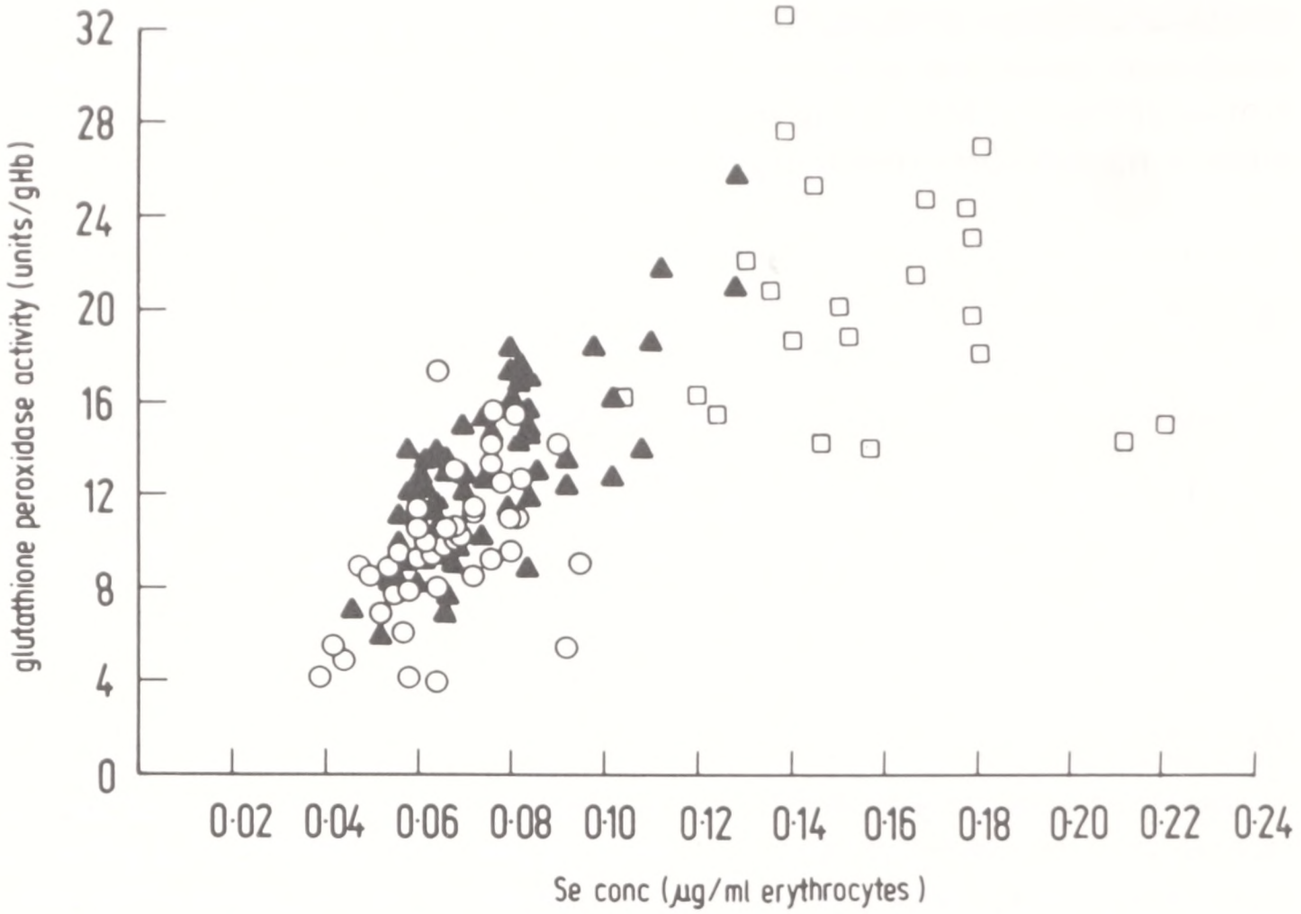


Figure 9–4. Relationship between selenium concentration of erythrocytes and glutathione peroxidase activity of whole blood for Otago patients (○); Otago blood donors (▲); and overseas subjects (□). (From Robinson and Thomson, 1981; reproduced with permission.)

Methods

Several accurate technical methods exist for the measurement of selenium in biologic tissues, including fluorometric analysis, atomic absorption spectrophotometry, and neutron activation analysis. This last technique has the advantage that it does not destroy the specimen, which can be reused for other analyses. The coefficient of variation for duplicate selenium neutron activation analysis in serum is 7.5 percent, and for nails approximately 4 percent (J. S. Morris, personal communication). Automated methods are available for the assay of glutathione peroxidase (McAdam et al., 1984).

Storage

As an inorganic element, selenium is not prone to degradation in storage. Levander (1985) states that glutathione peroxidase, however, loses activity in stored samples, although no quantitative results are described.

Measurement of Other Determinants

Lloyd and co-workers (1983) reported that plasma, whole blood, and erythrocyte selenium, but not whole blood glutathione peroxidase, were reduced in older

persons. Hunter and colleagues (1987) similarly observed an inverse relation between age and toenail selenium. Smoking is associated with reduced selenium status. Among men older than 30 years, plasma, whole blood, and erythrocyte selenium, and glutathione peroxidase activity, were all significantly reduced among smokers compared with nonsmokers (Lloyd et al., 1983). Toenail selenium was significantly lower among female smokers than nonsmokers, and an inverse dose-response relation was observed between increasing daily cigarette consumption and decreasing toenail selenium levels (Hunter er al., 1987). Serum selenium is reduced among alcoholics (Korpela et al., 1985), perhaps because alcoholic beverages are low in selenium, and thus, dietary selenium is reduced in those who consume a high proportion of total calories from alcohol. The effect of moderate alcohol consumption on selenium status is minimal (Hunter et al., 1987).

The relationships of selenium with disease may be modified by other antioxidants. Willett and colleagues (1983c) observed, for instance, that the relative risk of cancer among subjects with low serum selenium was increased among those who also had low serum vitamin E. Thus, information on vitamin E, and possibly vitamin C status, may be useful for the interpretation of epidemiologic studies involving selenium.

IRON

Iron deficiency may be the most common deficiency of a single nutrient in both the developing and the developed world. Iron is essential for the formation of hemoglobin and many intracellular heme-enzymes. Dietary iron requirements are determined by blood loss (including menstrual blood loss), as well as by the needs of growth in children, adolescents, and pregnant women.

Intestinal iron absorption is closely regulated and inversely related to body iron stores. Reduced body iron stores may be caused either by inadequate dietary iron, excessive blood loss, or both. Thus, information on the extent and source of blood loss are essential to the interpretation of biochemical measures of iron status as markers of dietary intake. In many developing countries, blood loss due to hookworm infestation, or the demands of repeated cycles of pregnancy and lactation, combine with poor nutritional intake to produce biochemical and clinical evidence of iron deficiency. In addition, it is important to recognize that the bioavailability of iron from different foods varies considerably. Nonheme iron (mainly from vegetables) and heme iron (from meat and fish) have independent absorption pathways in the intestinal mucosa, heme iron being much more efficiently absorbed. Absorption of nonheme iron is influenced by other dietary factors. In a study of 49 subjects, (Hallberg and Rossander, 1984), consumption of 65 g of ascorbic acid in cauliflower more than tripled nonheme iron absorption, whereas 1 g of citric acid reduced it by two-thirds. Biochemical data may be used to identify groups whose iron requirements are not being met, but are a poor indication of dietary intake for individuals without substantial additional information about other dietary components, growth requirements, and blood loss.

Measures

Serum Ferritin

Ferritin is the principal iron storage protein, and serum ferritin provides the best single indicator of iron stores (Cook and Skikne, 1982). Ferritin is decreased in proportion to the frequency of phlebotomy, and increased in proportion to supplementation. Serum ferritin is increased in iron overload (Cavill et al., 1986). The average within-subject CV% for serum ferritin drawn from 13 adults over 15 days was 14.5 percent (for a technique with a within-assay CV% of 4.3%) (Pilon et al., 1981). In a study of healthy child-bearing women, Soustre and coworkers (1986) did not observe a significant correlation between total daily iron intake and serum ferritin. Meat intake, however, was a significant predictor of serum ferritin, suggesting that serum ferritin does provide a measure of bioavailable (heme) iron.

Serum Iron

Serum iron is very variable, and changes of more than 20 percent have been observed within 10 minutes, among healthy subjects (Cavill et al., 1986). This short-term variability makes serum iron an unreliable long-term measure.

Transferrin Saturation and Total Iron-Binding Capacity

Transferrin is the plasma iron-transport protein. Transferrin saturation is measured as the ratio of serum iron and total iron-binding capacity (TIBC). As the TIBC of the transferrin pool is relatively stable, transferrin saturation is determined principally by changes in the serum iron concentration and is, thus, equally subject to short-term variation.

Erythrocyte Protoporphyrin

A deficiency of iron results in impaired heme synthesis and the accumulation of protoporphyrin, a heme precursor, in erythrocytes. Measurement of erythrocyte protoporphyrin is a sensitive indicator of iron deficiency, but provides no information about iron overload. Cavill and colleagues (1986) have emphasized the advantages of this test for large-scale surveys: the small sample size (about 20 μl of blood) required, and the relative simplicity and reproducibility of the measurement.

Mean Corpuscular Volume

Microcytosis of erythrocytes is a morphologic indicator of iron deficiency, but can also be caused by thalassemia and inflammation.

Hemoglobin or Hematocrit

More severe and long-standing iron deficiency results in reduced hemoglobin. Overhydration, the hemoglobinopathies, vitamin B_{12} deficiency, and a wide variety of chronic disease conditions also cause lowered hemoglobin.

Hair and Nails

The biologic significance of iron in hair and nails remains to be determined.

Other Determinants

The major determinant of iron stores, other than dietary intake, is blood loss. A history of blood donations, and a menstrual history in women, are valuable in the interpretation of iron measurements. Serum ferritin is significantly increased among oral contraceptive users, possibly due to reduced menstrual blood loss (Frassinelli-Gunderson et al., 1985).

Summary

Pilch and Senti (1984a) have emphasized that the use of several biochemical indicators of iron status provides a more sensitive and specific assessment than any single indicator alone. If the aim is to categorize subjects according to diagnostic groups, then the choice of appropriate cutoff values is a source of continuing controversy. Extensive additional information about the other determinants of iron stores is necessary to interpret biochemical measures of iron status in individuals.

SODIUM

On the basis of geographic correlation studies, it has been suggested that salt intake is positively correlated with blood pressure; however, results from other observational study designs and experimental studies have failed to confirm this speculation. Consumption of salt-cured foods has been linked with cancer of the esophagus and the stomach (Committee on Diet, Nutrition, and Cancer, 1982). A continuing problem in studies of sodium intake has been its measurement, because day-to-day variation in sodium intake is high (Caggiula et al., 1985).

Measures

Measurement of sodium in the blood provides almost no information about sodium intake, due to tight homeostatic control mechanisms that minimize variation in blood sodium. Urinary sodium, in contrast, is a good measure of short-term sodium intake. In a study of free-living subjects, average absorption of sodium from food was 98 percent, and average urinary excretion was 86 percent (Holbrook et al., 1984). A measurement of average urinary output is thus a good indicator of dietary intake. The problem is that sodium excretion is determined by recent intake, and is thus almost as variable as intake itself. Liu and co-workers (1979a) estimated the ratio of day-to-day within-person to between-person variation in 24-hour sodium excretion to be greater than three. Thus, a single 24-hour urine collection is a poor guide to true long-term intake. Nevertheless, Caggiula and colleagues (1985) observed a high correlation ($r = 0.61$) between sodium estimated from a 6-day food record and a single 24-hour urine measurement among 50 adult subjects. Similarly, Holbrook and colleagues (1984) reported a correlation of $r = 0.76$ between sodium intake calculated from analysis of four 7-day duplicate meal collections and four 24-hour urine sodium

measurements among 28 individuals. Although these correlations are impressive, it is important to note that the dietary reference periods were relatively short, and that correlations with long-term intake would be substantially lower. It seems probable that several 24-hour urinary sodium measurements collected over a number of months would provide a reasonable estimate of long-term sodium intake (Beaton and Chery, 1986).

Given the logistic difficulties that attend 24-hour urine collection in large studies, the alternative of using overnight specimens has been considered. Liu and colleagues (1979b) estimated a correlation coefficient of 0.73 between mean 24-hour and overnight urine sodium excretion measurements in children. Watson and Langford (1970) observed a correlation of $r = 0.76$ between overnight and 24-hour sodium excretion measurements among adults. Given the relative ease of obtaining overnight urine specimens, it appears that for many epidemiologic studies, collection of a greater number of overnight urine samples per subject may be preferable to a smaller number of 24-hour collections. If an absolute, rather than relative, estimate of 24-hour sodium excretion is required, then a 24-hour sample is necessary due to the diurnal variation in sodium excretion. Of interest is the report by Farleigh and co-workers (1985) who observed a statistically significant correlation ($r = 0.28$) between the sodium concentration of saliva, and the preceding (but not the same) day 24-hour urine sodium excretion.

Storage

As long as the total original volume is noted, preservation of an aliquot of urine results in substantial saving of freezer space.

Stability

Sodium is stable in frozen urine.

Measurement

Sodium can be accurately measured in urine by flame atomic absorption spectrophotometry or ion-selective electrode potentiometry.

Measurement of Other Determinants

Holbrook and colleagues (1984) observed a small, but significant, decrease in the percentage of dietary sodium excreted in summer compared to winter, possibly due to increased excretion in sweat.

POTASSIUM

The principal epidemiologic interest in potassium relates to the putative association of low potassium intake with the development of hypertension

(McCarron et al., 1984), and, possibly, strokes (Khaw and Barrett-Connor, 1987).

Measures

Blood potassium is under tight homeostatic control and does not reflect potassium intake over its normal range. The proprotion of ingested potassium excreted in urine, 77 percent in the study of Holbrook and colleagues (1984), is less than that for sodium as relatively more potassium is excreted in stool. Nonetheless, the correlation between potassium intake and 24-hour urinary excretion was $r = 0.82$, and the correlation of a single 24-hour urine and 6 days of food records was $r = 0.62$ (Caggiula et al., 1985).

Measurement of Other Determinants

Potassium-wasting diuretics or potassium supplements may affect potassium excretion.

CALCIUM

It has been suggested that dietary calcium protects against hypertension onset (McCarron et al., 1984) and colorectal cancer (Garland et al., 1985). In addition, the relationship between dietary calcium and osteoporosis is a subject of much activity and considerable controversy.

Measures

Blood calcium is under tight homeostatic control and thus not a useful indicator of intake. Few data are available on the relation between calcium intake and urinary calcium. Twenty-four-hour calcium excretion was lower in subjects consuming a low-calcium diet whose calcium intake was estimated to be approximately 75 percent of those consuming a high-calcium diet (Castenmiller et al., 1985). Interestingly, in this study, sodium supplementation appeared to increase calcium excretion, independent of calcium intake. Among 22 omnivorous, healthy, male subjects, a calcium supplement of 1.12 g daily (a mean increase of 92 percent above normal dietary intake) resulted in a 20 percent increase in mean urinary calcium excretion (Aalberts et al., 1988). Although urinary calcium reflects intake to some extent, further information is needed about the nature of this relation before urinary calcium could be recommended as a measure of dietary calcium.

MAGNESIUM

Relative deficiency of dietary magnesium has been associated with hypertension in some studies (Joffres et al., 1987) and coronary heart disease (Luoma et al., 1983).

Measures

Serum magnesium is tightly controlled. One mechanism of magnesium regulation is urinary excretion, and a reduction in magnesium intake causes a reduction in urinary excretion (Wacker, 1980); however, the utility of urinary estimates of intake remain to be determined.

COPPER

Although copper-dependent enzymes and copper-containing proteins are numerous, clinical diagnosis of copper deficiency appears to be rare, presumably due to the wide distribution of copper in foods and cooking utensils. Concern has been expressed, however, that per capita copper intakes in the United States may be substantially lower than the recommended dietary allowance of 2 to 3 mg/day suggested by the Food and Nutrition Board of the National Research Council (Klevay et al., 1980). Reduced copper status has been associated with ischemic heart disease (Klevay, 1983) and cardiac arrhythmias (Reiser et al., 1985).

Measures

Serum and Plasma

Few data are available on the relation between dietary copper intake and serum or plasma levels. Plasma levels in patients on total parenteral nutrition solutions deficient in copper were observed to fall steadily, until copper repletion restored normal levels (Solomons et al., 1976). No increase in serum copper was observed, however, among seven subjects who consumed 10 mg of copper per day for 12 weeks (Pratt et al., 1985). A further reason circulating copper levels are a poor guide to copper status is the large number of life-style factors (e.g., oral contraceptive use, smoking), and pathologic conditions (e.g., infections and inflammation), which profoundly alter blood copper concentration (Solomons, 1985).

Ceruloplasmin

Delves (1976) found 94 percent of circulating copper bound to the cuproprotein ceruloplasmin, and it has been suggested that this would be a more reliable measure of copper status. Ceruloplasmin production is also, however, influenced by many factors other than dietary intake (Solomons, 1985).

Erythrocyte Superoxidase Dismutase

Measurement of the function of the copper- and zinc-dependent enzyme, superoxidase dismutase (SoD), may be a better index of copper status. Among four men fed a copper-deficient diet for 4 months, erythrocyte SoD activity declined markedly in all four, whereas plasma copper declined in only one, and copper repletion restored normal SoD values (Milne et al., 1982). In a similar

experiment on nine men consuming a low copper diet containing 20 percent fructose, erythrocyte SoD activity, but not plasma copper or ceruloplasmin levels, reflected both copper depletion and repletion (Reiser et al., 1985).

Hair

Copper can be measured in hair, but it is not elevated in Wilson's disease (Martin, 1964). Hambidge (1973) demonstrated a progressive increase in mean copper concentration in hair with increasing distance from the scalp, data highly suggestive of external contamination.

Nails

Copper concentration was moderately elevated in the fingernails of two out of three patients with Wilson's disease (Martin, 1964). It remains to be determined, however, whether nail levels reflect normal dietary intake.

Urine

The major excretory route for copper is bile, and only small amounts are lost in urine (Cousins, 1985). In a study of 12 subjects who collected 24-hour urine specimens for 5 days, the between-day CV% was 53 percent, suggesting that large day-to-day fluctuations in urinary copper excretion make this an unreliable indicator of long-term copper status (Yang et al., 1986).

Methods

Atomic absorption spectrophotometry (AAS) is the most common analytic technique for copper in serum or plasma. There are several methods available for measurement of ceruloplasmin and SoD activity; however, these are all quite labor-intensive (Klasing and Pilch, 1985). Hair and nail copper can be measured using neutron activation analysis, or AAS.

Storage

Copper is stable in blood; the stability of ceruloplasmin and SoD is uncertain.

Measurement of Other Determinants

Blood levels of copper are sensitive to a number of nondietary determinants. An example of the extent of this problem is the fact that the reference range for serum copper among oral contraceptive users is higher than, and does not even overlap, the range for nonusers (Alpers et al., 1983). Several other dietary components appear to influence copper absorption; amino acids appear to increase absorption, whereas ascorbic acid, fiber, and zinc have been observed to decrease copper absorption (Cousins, 1985). Collection of information about all of these factors, and in particular, the use of vitamin C and zinc supplements, should theoretically aid in the interpretation of measures of copper status.

ZINC

Marginal zinc nutriture has been linked to a variety of clinical conditions as well as to suboptimal physiologic performance (Pilch and Senti, 1984b). No single measure of zinc status is a reliable measure of intake. Although zinc supplementation increases plasma zinc levels (Fischer et al., 1984), dietary zinc deprivation may not result in lowered plasma zinc (Prasad, 1978). Thus, zinc deficiency may coexist with a normal plasma zinc level (Sandstead et al., 1982). In a study of 24 elderly subjects, zinc intake was not significantly correlated with zinc concentrations in plasma, whole blood, or leukocytes (Bunker et al., 1984). Furthermore, blood zinc is sensitive to a wide range of nondietary factors (Solomons, 1985). Measurements of zinc in erythrocytes and hair have not been widely accepted as indices of dietary intake (Klasing and Pilch, 1985). No increase in hair zinc levels was observed among 52 pregnant women whose mean dietary zinc intake was estimated to be about 50 percent of the recommended daily allowance, and who consumed 20 mg of zinc daily for a mean of 4.5 months (Hunt et al., 1983). Among 13 men consuming a zinc supplement of 40 mg/day (almost four times their estimated dietary intake) for 12 weeks, a transient rise in serum zinc was observed, but levels declined to baseline after 4 weeks (Black et al., 1988). Suggested functional tests, such as taste and smell acuity and neuropsychologic function, are clearly not specific for zinc deficiency.

LIPIDS

The relation between diet and coronary heart disease has been of major interest over the last three decades, and the lipid composition of diet has been identified as a predictor of this disease (Shekelle et al., 1981; Kushi et al., 1985). The general acceptance of this association has been delayed by the controversy over the relation between the lipid composition of diet and the levels of cholesterol and lipoproteins in blood, cells, and fat. In recent years these relationships have become better understood, principally because many sources of variance in both dietary fats and biochemical parameters have been determined (see Chapter 15). These developments allow us to better evaluate the use of lipid biochemical indicators as measures of lipid intake.

Given the epidemiologic interest in the relation of total fat intake to the occurrence of both cardiovascular disease and cancer, it is unfortunate that there is no biochemical measure of total dietary fat. This also leads to problems in the interpretation of the measurement of specific lipids. A finding, for instance, that the concentration of a specific fatty acid is equal in the adipose tissue of cases of a disease and of controls, could have two interpretations. The absolute intake of the fatty acid might be equal in both groups in the context of equal total fat intake. On the other hand, the absolute intakes may differ, but differences in the same direction in total fat intake may mean that consumption of the fatty acid as a percent of total fat is equal and there is no relative difference. These alternatives would lead to different causal interpretations and dietary recommenda-

tions, but without a measure of total fat intake it is not possible to distinguish between them.

Lipid Physiology: A Simple Overview

Dietary fats are largely broken down in the small intestine into cholesterol and triglycerides (molecules of glycerol and fatty acids). Cholesterol, derived solely from animal sources, is emulsified with bile acids and absorbed into the portal circulation. Triglycerides, available from both plant and animal sources, are hydrolyzed into monoglycerides and diglycerides by pancreatic lipase, before absorption through the intestinal mucosa, reformation into triglycerides, and transport via blood and lymph in chylomicrons (Linscheer and Vergroesen, 1988). Dietary cholesterol is transported to the liver, which also synthesizes cholesterol de novo. Endogenous cholesterol synthesis in the liver is more than twice the typical American dietary cholesterol consumption, and feedback mechanisms cause changes in endogenous synthesis to partially counteract changes in dietary intake (Samuel et al., 1983). Cholesterol and triglycerides are synthesized in the hepatocytes, complexed with proteins and phospholipids to form lipoproteins, and then secreted into the blood. There are four basic types of lipoproteins: chylomicrons, very low-density lipoproteins (VLDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL), listed in order of decreasing size and increasing density. The protein moieties of these transport particles are called apolipoproteins, some of which function as activators of the enzymes of cholesterol and triglyceride metabolism and as ligands for cell surface receptors. VLDL and chylomicrons are the primary transport particles for triglycerides to cells, whereas LDL is the major vehicle for cholesterol transport. HDL scavenges excess cholesterol from cells, returning it to other lipoproteins, to the liver, and to extrahepatic cells that require cholesterol for steroid hormone production and membrane biosynthesis.

There are large differences in plasma total cholesterol and LDL-cholesterol (LDL-C) levels between populations, and the diversity within most populations is large. Population differences in average levels of LDL-cholesterol are determined by differential rates of LDL production and catabolism. It is not known, however, if these rates are controlled by dietary intakes or other environmental factors (International Collaborative Study Group, 1986). Within populations, large interindividual differences exist in the ability of subjects to suppress endogenous cholesterol synthesis when challenged with increased dietary cholesterol (McNamara et al., 1987). Large differences have also been observed in individual responses of plasma concentrations of total cholesterol and LDL-C to the substitution of saturated fatty acids for unsaturated fatty acids (Grundy and Vega, 1988). These differences, presumably genetic in origin, may be the most important determinants of individual cholesterol levels.

SERUM TOTAL, LDL-CHOLESTEROL, AND HDL-CHOLESTEROL

Both serum total cholesterol and LDL-C have a positive relation with coronary heart disease incidence, whereas HDL-cholesterol (HDL-C) is inversely related

(Gordon et al., 1977; Wallace and Anderson, 1987). The relation between dietary cholesterol intake and these measures has, however, been controversial. The primary argument has been between those who claim that correlations exists and those who claim they are absent. Among adherents to the former view, there has been much discussion about the extent and shape of the relation. Based on metabolic ward studies, both Keys and colleagues (1965) and Hegsted and colleagues (1965) proposed equations that predict serum cholesterol from dietary intake of cholesterol, saturated, and polyunsaturated fatty acids. The Keys equation* proposes that plasma cholesterol increases in proportion to the square root of dietary cholesterol intake (in mg/1000 kcal per day). It thus predicts that plasma cholesterol will increase more for a given increase in cholesterol consumption if this increase occurs from a low, rather than high, baseline. The Hegsted equation† proposed a linear relationship. Both equations appear to perform reasonably well over the usual range of dietary cholesterol in developed countries. In the low and high range of cholesterol intake, however, the nonlinear Keys equation appears to be superior (Keys, 1984). The fact that dietary intake of saturated and polyunsaturated fatty acids are predictors of serum cholesterol indicates that dietary cholesterol is not the only dietary determinant of serum levels. The performance of plasma cholesterol as an indicator of dietary intake can be partially deduced from the shape of the Keys relation. Plasma cholesterol rises less steeply as dietary cholesterol intake increases. Thus, it will probably be a better index of cholesterol consumption in the low, rather than in the moderate and high range. The increases predicted by these equations are small in comparison with the known variance of serum cholesterol in the population. This suggests that although they may predict changes in serum cholesterol in response to dietary change for groups, the use of serum cholesterol to calculate the dietary cholesterol intake of individuals results in severe misclassification (see Chapter 1).

Data from the many intervention studies and the few observational studies that have been conducted to study this issue support this deduction. Subjects in the Leiden Intervention Study (Arntzenius et al., 1985), whose diets were already low in cholesterol and whose dietary cholesterol was reduced from a daily mean of 89.6 to 29.5 mg/1000 kcal, were observed to have a reduction of serum cholesterol from 6.9 to 6.2 mmol/L. Thus, a dietary intervention that effectively reduced dietary cholesterol by two-thirds resulted in a 10 percent reduction in serum cholesterol. In the National Diet–Heart Study, among subjects whose dietary cholesterol intake was representative of normal U.S. intake (approximately 200 mg/1000 kcal per day), consuming a diet that reduced cholesterol alone, a reduction of approximately 100 mg/1000 kcal per day (approximately a 50% decrease in cholesterol intake) was associated with a less than 3 percent reduction in serum cholesterol (National Diet Heart Study Research

*The Keys equation is $\Delta y = 1.35\ (2\Delta S - \Delta P) + 1.5\Delta Z$ where Δy = change in serum cholesterol (mg/dl), ΔS and ΔP = change in dietary intake of saturated and polyunsaturated fatty acids expressed as percentage of calories, $\Delta Z = (x_1^{0.5} - x_2^{0.5})$ where x_1 and x_2 are the dietary cholesterol of the two diets being compared in mg/1000 kcal.

†The Hegsted equation is $\Delta y = 2.16\Delta S - 1.65\Delta P + 0.176\Delta C$, where ΔC is change in cholesterol intake in mg/1000 kcal.

Group, 1968). In an intervention study in the Netherlands (Katan et al., 1986), 94 healthy men and women had their cholesterol intake increased from a daily average of 49 to 234 mg/1000 kcal; serum cholesterol increased from an average of 5.25 mmol/liter by an average of 0.50 mmol/L after 13 days. In this study, two groups of 15 "hyporesponders" and 17 "hyperresponders" were identified, and the hyporesponders had a significantly lower increase in serum cholesterol than the hyperresponders, when challenged with a larger increase in dietary cholesterol (349 mg/1000 kcal/day), indicating that variability of response to dietary cholesterol exists within populations. In cross-sectional observational studies, correlations between dietary cholesterol intake and serum cholesterol values depend on the range of cholesterol intake in the population being examined. In a macrobiotic vegetarian population with a mean dietary cholesterol intake of 42.3 mg/1000 kcal per day, the unadjusted correlation coefficient between dietary cholesterol consumption and serum cholesterol was $r = 0.46$ (Kushi et al., 1988). Shekelle and co-workers (1981), however, observed an unadjusted correlation of $r = 0.078$ between the calculated Keys score (which includes cholesterol, saturated and polyunsaturated fatty acid intake) and serum cholesterol among 1900 middle-aged men whose mean dietary cholesterol intake was 240.5 mg/1000 kcal per day. It is thus apparent that although serum cholesterol reflects dietary intake at low intake levels, it is a poor measure of differences in cholesterol intake between individuals consuming typical Western diets.

About two-thirds of total serum cholesterol is contained in LDLs, and the relation between cholesterol intake and level of LDL is very similar to the relation with total cholesterol. In a large dietary intervention study among hypercholesterolemic men (Glueck et al., 1986), decreased cholesterol intake was associated with reductions in both total plasma cholesterol and LDL-C. Dietary saturated fat (a positive relation) and dietary polyunsaturated fat (a negative relation), however, were stronger determinants of total and HDL-C. Numerous studies have demonstrated that replacement of dietary saturated fatty acids with polyunsaturated fatty acids lowers LDL-C and total cholesterol, although the mechanism is still unclear (Beynen and Katan, 1985b). Not all dietary saturated fatty acids, however, have the same metabolic effects. A diet high in stearic acid (18:0) did not raise total cholesterol or LDL-C levels among 11 men, in contrast with increases seen with a diet high in palmitic acid (16:0) (Bonanome and Grundy, 1988). Differences in HDL-C between population groups with varying dietary fat intakes are less than those seen for LDL-C, and changes in HDL-C observed in intervention studies are generally less than those observed for LDL-C. This suggests that HDL-C levels are less sensitive to dietary fat intake than LDL-C levels (Sacks et al., 1985). In summary, lipoproteins may be useful in intervention studies as markers of response. In observational studies, however, lipoprotein levels, like serum total cholesterol, are poor indicators of dietary cholesterol or fatty acid intake, at least in developed countries.

Laboratory Techniques

A variety of reliable, automated chemical and enzymatic methods are available for the measurement of total cholesterol in serum. Lipoproteins may be mea-

sured by ultracentrifugation, which is the generally accepted reference method, or by column chromatography, electrophoresis, or precipitation reactions (Naito and David, 1984). Use of certified reference materials, such as those available from the Centers for Disease Control, is essential to ensure the validity of total cholesterol and HDL-C measurements (Hainline et al., 1978).

Measurements of cholesterol and the lipoproteins exhibit an appreciable degree of within-person day-to-day variation. Coefficients of variation between the initial and second visit for 7055 white participants in the Lipid Research Clinics Prevalence Study were 8 percent for fasting plasma cholesterol and 25 percent for fasting plasma triglycerides (Jacobs and Barrett-Connor, 1982). Mjos and co-workers (1979) observed within-person coefficients of variation of approximately 8 percent for total cholesterol, 9 percent for HDL-C, 11 percent for LDL-C, and 26 percent for triglycerides. Even under metabolic ward conditions with constant dietary intake, coefficients of variation of 3.1 to 9.0 percent for serum cholesterol have been observed (Hegsted and Nicolosi, 1987). This degree of variability, the cause of which is poorly understood, means that a single, or even several, measurements of serum cholesterol, lipoproteins, and triglycerides, will misclassify many subjects with respect to their long-term average levels. This further reduces our ability to observe any underlying association with dietary intake.

On a technical note, increases of approximately 10 percent in the levels of total cholesterol, HDL-C, LDL-C, and VLDL-C, and triglycerides, were observed when blood was taken after 30 minutes of standing relative to 30 minutes in the supine position (Kjeldsen et al., 1983). These changes, attributable to an orthostatic decrease in plasma volume, emphasize the necessity for strict standardization of procedures for blood drawing and subject posture in the measurement of all biochemical parameters in blood.

Stability

Total cholesterol is stable in frozen serum (Wood et al., 1980). Repeated freezing and thawing degrades lipoproteins, with HDL being particularly susceptible. HDL-cholesterol values were more stable when stored at −60°C than at −20°C over 10 months (Cooper, 1979). Thus, storage at the lower temperature seems advisable; the validity of measurements made in serum or plasma stored for more than a year needs to be determined.

Measurements of Other Determinants

A further reason that cholesterol and lipoproteins are poor indicators of dietary intake is the influence of nondietary factors. Levels of LDL-C are positively correlated with Quetelet's index and smoking (Glueck et al., 1986). HDL-cholesterol is increased by exercise and alcohol consumption, and has an inverse relation with smoking and body mass index (Hulley et al., 1977). Even after adjusting for these factors, the correlation of lipoproteins with dietary cholesterol remains poor.

FATTY ACIDS

Specific fatty acid levels in blood, cell membranes, and subcutaneous fats are more promising indicators of dietary fat intake than cholesterol or lipoprotein measurements. Although this field is not new, recent improvements in gas chromatography and HPLC methods have made the measurement of many individual fatty acids and their isomers more feasible in large studies.

Fatty acids are numbered by their carbon chain length. Saturated fatty acids have no double bonds between carbon atoms, monounsaturated fatty acids have one, and polyunsaturated fatty acids have two or more double bonds. The standard numbering system for fatty acids gives the number of carbon atoms, the number of double bonds (after a colon) and the position of the first double bond (after the Greek letter ω) counting from the end of the carbon chain opposite the carboxyl group. Thus, oleic acid (18:1ω9) is a monounsaturated fatty acid with its single double bond nine carbon atoms from the C18 end of the chain. In *trans* isomers of fatty acids, the two hydrogen atoms attached to the carbon atoms linked by a double bond are on opposite sides of the molecule. These hydrogen atoms are in the *cis* configuration, on the same side of the molecule, in most dietary unsaturated fatty acids.

Much work remains to determine the relations between diet and individual chromatographic peaks representing individual fatty acids or groups of peaks representing common metabolic pathways. It seems reasonable to expect that the best markers of dietary intake would be the fatty acids that cannot be endogenously synthesized from carbohydrates. These are largely the ω-3 fatty acids (medium chain lengths derived from some plants, and long chain lengths from sea animals), the ω-6 family (mostly from vegetables,) *trans* fatty acids (primarily from processed fats and ruminants), and odd-number and branched chain fatty acids (from dairy products). Linoleic acid (18:2ω6) is the principal dietary essential fatty acid, and it is capable of being metabolized to longer-chain, more highly unsaturated, ω-6 fatty acids, including arachidonic acid (20:4ω6). Similarly, linolenic acid (18:3ω3) gives rise to a sequence of ω-3 fatty acids. Oleic acid (18:1ω9) may be endogenously synthesized, and is the starting point for synthesis of ω-9 fatty acids. The relation between dietary intake of these three principal precursors (18:3ω3, 18:2ω6, and 18:1ω9) and the levels of their metabolites in human tissues is complex, largely because intake of one inhibits the elongation and unsaturation of the others (Holman, 1986). The majority of studies comparing fatty acid intake with fatty acid levels have concentrated on linoleic acid and oleic acid. More recently, limited information on the interpretation of other fatty acid peaks has become available.

Caution is appropriate in the interpretation of markers of fatty acid intake. Results are usually expressed as relative percentages of total fatty acid. Thus, an increase in dietary intake of one fatty acid, if incorporated into the substrate for analysis, results in a decrease in the relative amounts of all other fatty acids. This decrease should not necessarily be taken as evidence for a metabolic interaction. Such interactions may exist, however (e.g., the inhibition of metabolism of ω3

fatty acids by dietary intake of ω6 fatty acids described by Holman, 1986), and further complicate the interpretation of relative differences with respect to dietary intake. Metabolic inhibition of one fatty acid by another leads to the possibility that the between-person intake of a particular fatty acid may be confounded by differences in the between-person intake of an interacting fatty acid. Much work needs to be done to define the nature of these relations. In general, however, biochemical fatty acid profiles may be interpreted as indicating relative patterns of fatty acid intake, rather than absolute amounts.

Plasma Fractions, Cell Membranes, and Adipose Tissue

The measurement of individual fatty acids is further complicated by the fact that they may be measured in erythrocytes, platelets, and adipose tissue, as well as several lipid subfractions found in plasma. Data are sparse on the relations of fatty acid intake and their levels in these substrates, and interpretation of the existing data is frequently difficult because otherwise comparable studies have often measured fatty acid levels in different fractions.

Individual fatty acids may be measured in the cholesteryl ester, phospholipid, and triglyceride fractions of serum or plasma, or as free fatty acids. These fractions are usually separated by thin layer chromatography before proceeding with isolation and identification of individual fatty acids. Cross-sectional studies reveal major differences in the relative contribution of individual fatty acids to different fractions among subjects eating a Western diet. Linoleic acid concentration is typically approximately two to three times higher than oleic acid in the cholesteryl ester fraction of plasma. This predominance is reduced to a factor of between one and a half to two, in plasma phospholipids. The ratio is reversed in triglycerides, where oleic acid predominates. Oleic acid is also the major free fatty acid found in plasma, and is three to four times more common than linoleic acid in both platelet phospholipids and adipose tissue. Although palmitic acid levels are higher than stearic acid levels in almost all these substrates, the ratio of palmitic acid to stearic acid varies from about two in plasma phospholipids, to more than ten in plasma cholesteryl esters. Arachidonic acid (20:4ω6) constitutes 2.3 percent of total fatty acids in plasma triglycerides, 11.5 percent in plasma phospholipids, less than 1 percent in adipose tissue, but 27 percent in platelet phospholipids (Hirsch et al., 1960; Dayton et al., 1966; Phillips and Dodge, 1967; Manku et al., 1983; Sacks et al., 1987; Wood et al., 1987). The variation in fatty acid composition among plasma lipid fractions suggests that total plasma fatty acid levels is influenced by nondietary factors that determine the relative proportions of these fractions. Thus, measurements of fatty acid levels in specific lipid fractions should provide more precise indicators of dietary intake.

The distribution of fatty acids in the phospholipid fraction of plasma is closely related to their distribution in the phospholipids of erythrocyte membranes. In a study of 10 healthy subjects, Phillips and Dodge (1967) observed correlations between fatty acids in plasma and erythrocyte phospholipids of $r = 0.42$ (stearic acid), $r = 0.80$ (palmitic acid), $r = 0.51$ (oleic acid), and $r = 0.73$ (linoleic acid). Changes in the oleic and linoleic acid content of diet are reflected

in erythrocyte phospholipids within weeks of dietary change, suggesting that membrane fatty acids have a dynamic relation with plasma levels and are not fixed at the time of erythropoesis (Farquhar and Ahrens, 1963). In a study of 67 healthy men, Boberg and colleagues (1985) observed a correlation of $r = 0.78$ between the proportion of linoleic acid in platelet phospholipids compared with plasma phospholipids.

These differences in fatty acid composition of various tissues or plasma lipid fractions are due to the endogenous production of some fatty acids, the different roles of these fractions as vehicles for fatty acid transport, and the physiologic functions of individual fatty acids. In some tissues, the fatty acid levels may largely passively reflect diet, whereas in others these levels may be actively regulated, which is likely to be true when they are playing an important structural or functional role. The situation is further complicated if we examine the relation of dietary change and fatty acid levels, as factors such as the half-lives of cholesteryl ester fractions or red cell membranes are additional influences on the rate of incorporation of individual fatty acids. For use as an epidemiologic marker of fatty acid intake, the choice of a specific lipid fraction or substrate for measurement depends largely on two characteristics: the responsiveness or sensitivity of the substrate to changes in dietary fatty acids and the degree to which fatty acid levels in the substrate integrate intake over time. For example, in the study of Sacks and colleagues (1987), the largest responses to changes in intake of both linoleic and oleic acid were measured in the triglyceride fraction of plasma. Although this responsiveness is an advantage, the triglyceride fraction could be less useful as a measure of long-term intake if the day-to-day variation in fatty acid content was greater than for the other lipid fractions. Many more data are needed before recommendations can be made on which fatty acid measurements reflect dietary intake, and for each fatty acid, which plasma fraction or tissue is most suitable for measurement. If a generalization may be made at this stage, it is that the long half-life associated with fatty acid turnover in adipose tissue probably means that this tissue provides the best measure of long-term intake of the dietary fatty acids that are incorporated into fat. In studies seeking to monitor short-term dietary change, plasma lipid fractions, and erythrocyte or platelet membrane phospholipids may provide a more rapidly responding measure.

Linoleic, Linolenic and Oleic Acid

Cross-sectional studies of fatty acid intake and fatty acid profiles measured in plasma have generally failed to demonstrate strong positive relations. In a small group of 29 healthy free-living volunteers whose normal diet was unusually well-characterized, no consistent correlation between dietary linoleic acid intake and plasma linoleic acid concentration was observed (Reeves et al., 1984), possibly because the between-person variation in intake appears to have been small. In a geographic correlation study, where much larger dietary differences were involved, mean intakes of linoleic acid among subjects in Italy, Finland, and the United States correlated with the mean levels of linoleic acid in plasma (Dougherty et al., 1987). The daily estimated consumption of linoleic acid (as a per-

centage of energy intake in kcal) was 2.2 percent in the Italian center, 3.5 percent in the Finnish center, and 4.9 percent in the U.S. center. The corresponding percent contributions of linoleic acid to major fatty acids measured in the cholesteryl ester fraction of plasma were 33.4 percent, 41.2 percent, and 57.4 percent. In an intervention study in which 17 healthy adults increased their linoleic acid consumption from 5 to 15 percent of calories over 4 weeks, the linoleic acid content of plasma rose 33 percent in the triglyceride fraction, 7 percent in the cholesteryl ester fraction, and 8 percent among phospholipids (Sacks et al., 1987). Dayton and colleagues (1966) reported on 393 institutionalized men whose diet was altered to increase the linoleic acid content from 11 to almost 40 percent of total fatty acid while oleic acid intake was held constant. Linoleic acid measured in the cholesteryl ester, phospholipid, and triglyceride fractions of serum increased substantially, the largest change being seen in the serum triglyceride fraction (Fig. 9–5). This situation is analogous to that for cholesterol: as plasma linoleic acid can discriminate between groups with relatively large differences in intake, but performs less well on an individual basis.

For most purposes, the best readily available tissue for fatty acid measurement is adipose tissue. Hirsch and colleagues (1960) demonstrated that subcutaneous fat aspiration with a needle and syringe was a simple, virtually painless method of obtaining a fat sample. The fatty acid composition of subcutaneous fat is very similar to omental fat (Gurr et al., 1982). Several well-conducted studies have defined the kinetics of linoleic acid in adipose tissue. In the study of Dayton and co-workers (1966) (Fig. 9–6), the linoleic acid content of total fatty acid in subcutaneous fat increased from 11 to 32 percent after 5 years among elderly, institutionalized men whose diet was altered to increase the linoleic acid content from 11 to almost 40 percent of total fatty acid while oleic acid was held constant. Repeated sampling from these subjects allowed the half-life of adipose tissue linoleic acid to be estimated at about 680 days. Thus, the fatty acid content of subcutaneous fat provides a qualitative measure of fatty acid intake over the previous 2 or more years. Based on pooled data from seven studies, Beynen and co-workers (1980) observed a correlation of $r = 0.80$ between the percentage of polyunsaturated fatty acids relative to total fatty acid intake and the relative percentage of adipose tissue polyunsaturated fatty acids. These relations, observed in intervention and geographic correlation studies, have also been observed in cross-sectional studies conducted within single populations. In a study of 164 Scottish men, Thomson M. and colleagues (1985) observed a correlation of $r = 0.57$ between dietary linoleic acid intake determined from 7-day weighed diet records and the relative proportion of linoleic acid in adipose tissue. In a study of 59 Dutch women, van Staveren and colleagues (1986) also observed a high correlation ($r = 0.70$) between the dietary linoleic acid content of diet and subcutaneous fat. As this relation was observed despite a relatively small between-person coefficient of variation in the study group ($CV_b\% = 23\%$), this strong correlation is good evidence that the linoleic acid content of adipose tissue is a reliable indicator of dietary linoleic acid intake. Part of the reason for a stronger correlation between diet and adipose as compared with plasma levels may be that fatty acid concentrations are more regulated in the cholesteryl ester and plasma phospholipid fractions and thus respond less to alteration in intake.

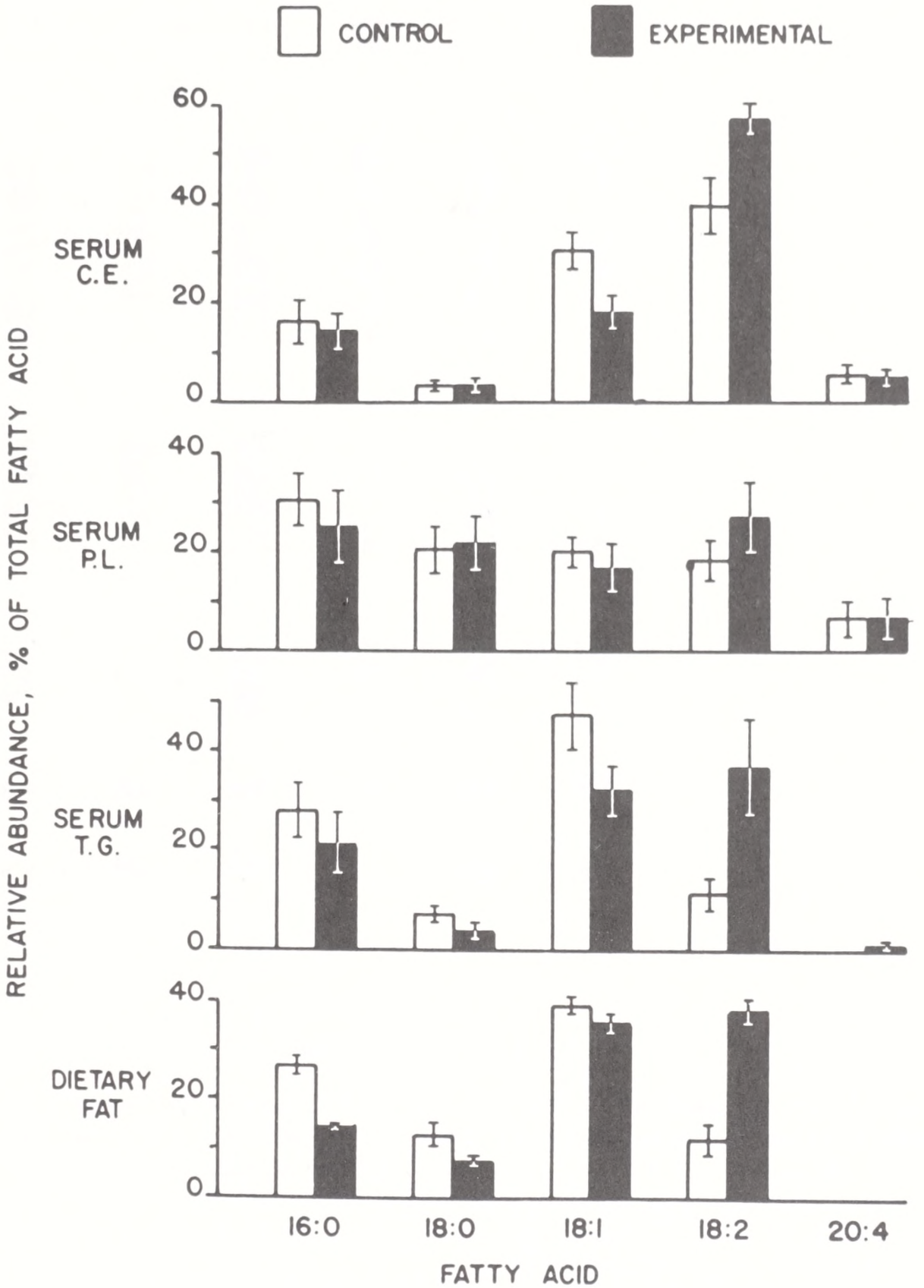


Figure 9–5. Fatty acids in the major serum lipid fractions (C.E. = cholesteryl ester, P.L. = phospholipids, T.G. = triglycerides) compared with dietary fat. Subjects were elderly, institutionalized men whose diet was unaltered (control), or whose diet was altered (experimental) to increase the linoleic acid content from 11 percent to almost 40 percent of total fatty acid while oleic acid was held constant. Data are presented for 10 subjects in each group after 3 years. Subjects were selected for adherence to dietary protocol of 88.6 percent or better. Values are shown as mean ±SD. Note the larger change in linoleic (18:2) acid and oleic (18:1) acid response in the triglyceride fraction, relative to the cholesteryl ester and phospholipid fractions. (From Dayton et al., 1966; reproduced with permission.)

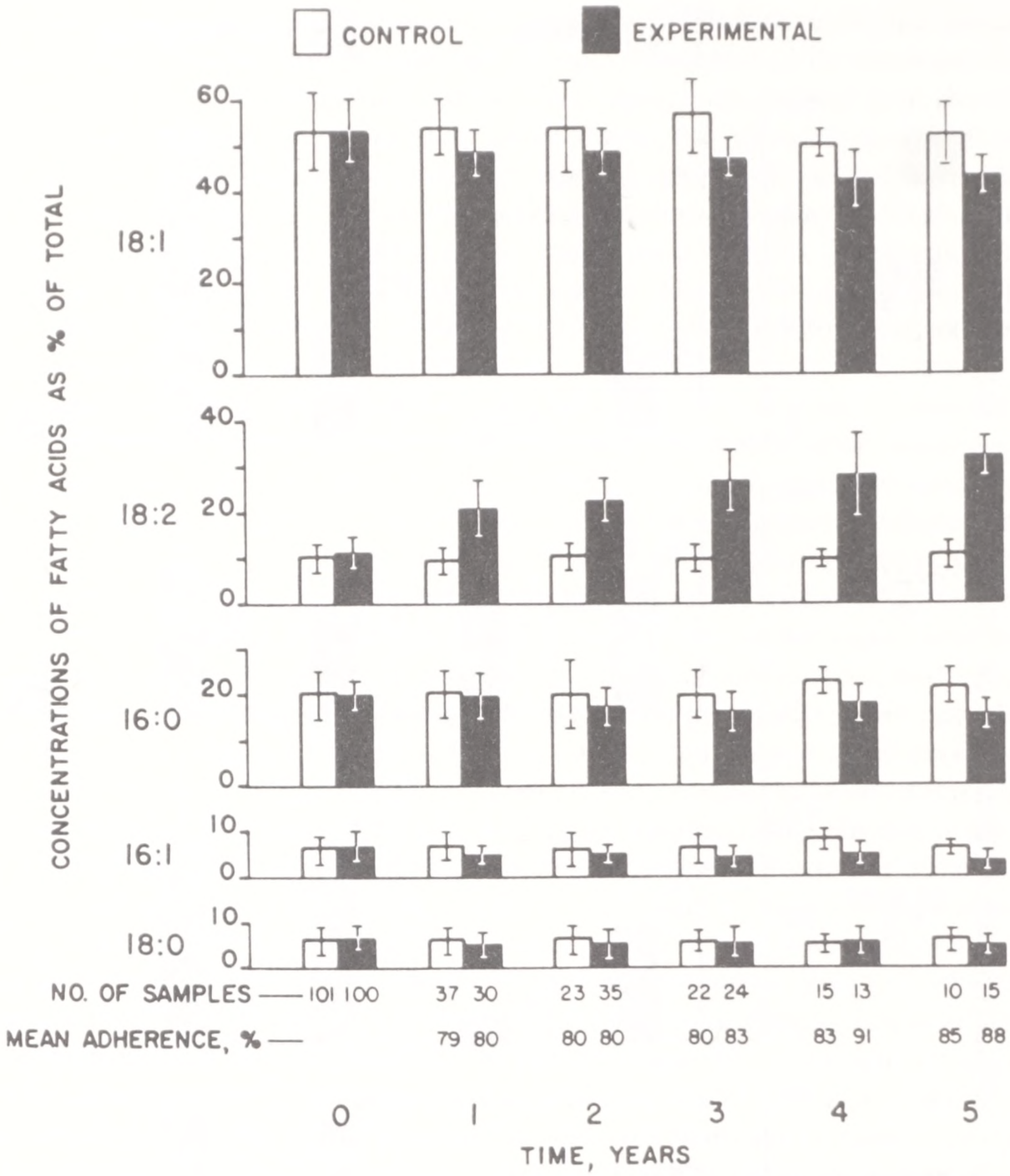


Figure 9–6. Fatty acids of aspirated adipose tissue, in control and experimental subjects. Subjects were elderly, institutionalized men, and men in the experimental group who had their diet altered to increase the linoleic acid content from 11 percent to almost 40 percent of total fatty acid while oleic acid was held constant. Time intervals shown are ±30 days. Values are shown as mean ±SD. (From Dayton et al., 1966; reproduced with permission.)

(Compare the response to an increase in dietary linoleic acid measured in serum cholesteryl ester, phospholipids, and adipose tissue in Figs. 9–5 and 9–6.)

Marine Polyunsaturated Fatty Acids

A substantial literature supports the idea that long-chain ω-3 marine fatty acids measured in human tissues do reflect dietary intake. Levels of eicosopentaenoic acid (20:5ω3) and docosahexaenoic acid (22:6ω3) measured in serum phospholipids were much higher in residents of a Japanese seaside fishing village than

among inland Americans (3.7 times higher for 20:5ω3; 9.9 times higher for 22:6ω3) (Yamori et al., 1985). Similar differences were noted in plasma and platelet eicosopentaenoic acid concentrations when Greenland Eskimos, having very high marine oil consumption, were compared with Danish controls (Dyerberg et al., 1975; Dyerberg and Bang, 1979). High fish oil consumption in the Japanese subjects was also reflected in a higher percentage of 20:5ω3 and 22:6ω3 in adipose tissue. It should be noted that the dietary differences involved here are large. The degree to which tissue fatty acids indicate relative marine oil intake in populations with a narrower range of consumption requires further investigation. After one subject consumed an "Eskimo diet" (seal, fish, and other marine animals) containing 14 g of eicosopentaenoic acid daily for 74 days, adipose tissue eicosopentaenoic acid concentration rose from 0 to only 0.5 percent. Eicosopentaenoic acid concentration in the phospholipids of erythrocytes, however, rose from 2.3 to 12.8 percent. After 36 months of consuming 1.8 g of eicosopentaenoic acid daily as fish oil, mean eicosopentaenoic acid concentration in adipose tissue among 28 subjects was 0.51 percent, whereas levels were not detectable among 28 controls (Sinclair and Gale, 1987). In a cross-sectional study, 15 subjects with a daily dietary ω-3 fatty acid intake of less than 0.10 g/day were observed to have 3.8 percent of ω-3 fatty acids in plasma phospholipids compared with 11 subjects consuming greater than 0.20 g/day whose mean plasma ω-3 fatty acid proportion was 5.2 percent ($p < 0.05$) (Silverman et al., 1988). These data suggest that eicosopentaenoic acid in adipose tissue, plasma, and erythrocytes discriminates between long-term fish eaters and nonfish eaters, however, the dose-response relation remains to be determined. Levels in platelet and erythrocyte membranes may be superior indicators of relative marine oil consumption.

Oleic Acid

Oleic acid is the principal monounsaturated fatty acid in the diet. Although most studies have demonstrated a positive relation between oleic acid intake and oleic acid measured in blood and adipose tissue, the relation is not as strong as that observed for linoleic acid, possibly because oleic acid can be synthesized endogenously from carbohydrate. In the study of Dougherty and co-workers (1987) referred to previously, the oleic acid contribution to major fatty acids in the cholesteryl ester fraction of plasma was 37.8 percent among Italian subjects and 18.6 percent among U.S. subjects, although the calculated oleic acid composition of the diets in the two centers was equivalent. The difference may be due to the fact that most of the oleic acid in the Italian diets is the *cis* isomer derived from olive oil, whereas much of that in the U.S. diet is probably the *trans* isomer derived from partial hydrogenation of vegetable oils in margarine. In the intervention study referred to earlier, an increase of oleic acid intake from 11 to 20 percent of calories had a variable effect on plasma oleic acid, increasing it by 14 percent among triglycerides, 0 percent in cholesteryl esters, and 9 percent in phospholipids (Sacks et al., 1987). In the study of Beynen and co-workers (1980), pooling seven data sets, the correlation between relative percentages of monounsaturated fats in diet and adipose tissue was $r = 0.46$. In a study in Crete,

high levels of oleic acid in adipose tissue were attributed to the diet high in olive oil (Fordyce et al., 1983). Thus, adipose tissue may reflect oleic acid intake reasonably well, however, more data are needed to define this relation.

Saturated Fatty Acids

The principal saturated fatty acids in both diet and tissues are palmitic acid (16:0), the most abundant, and stearic acid (18:0). Levels of these fatty acids in plasma do not provide a simple index of intake. In the study of Dougherty and co-workers (1987), for instance, although calculated intake of saturated fatty acids in the U.S. subjects (17.4% of calories) was double that of the Italian subjects (8.7% of calories), levels of saturated fatty acids in the cholesteryl ester fraction of plasma were actually lower in the U.S. subjects. It is probable that consumption of linoleic acid (higher in the U.S. subjects) and other polyunsaturated fatty acids reduces saturated fatty acid levels in plasma, which may thus be a better reflection of the dietary polyunsaturated to saturated fatty acid ratio. Among 20 healthy subjects, correlations between normal intake of total saturated fatty acids and fatty acid composition of triglycerides in adipose tissue were $r = 0.57$ for palmitic acid and $r = 0.56$ for stearic acid (Field et al., 1985). In the intervention study of Dayton and co-workers (1966), in which dietary intake of saturated fat (initially about 40% of total fatty acid) was approximately halved, differences in serum palmitic and stearic acid composition were slight after 3 years of the intervention, whether measured in the cholesteryl ester, phospholipid, or triglyceride fractions. Substantial reductions in palmitic and stearic acid levels in adipose tissue were observed, however, suggesting that long-term saturated fatty acid intake may be reflected in adipose tissue levels.

A novel approach was taken by McMurchie and colleagues (1984a), who compared fatty acid levels in human cheek epithelial cell phospholipids collected from vegetarians and a group of nonvegetarians. Cheek cells are collected in a noninvasive manner by asking subjects to rinse their mouths with distilled water, and centrifuging the expectorated fluid. Saturated fat as a percent of total fat intake (21% among vegetarians, 27% among nonvegetarians) was closely reflected in the saturated fatty acid composition of the cheek cell phospholipids (21% among vegetarians, 34% among nonvegetarians). Little difference in polyunsaturated fatty acid intake was present, and the monounsaturated fatty acid composition of phospholipids was actually higher among the vegetarians despite a substantially lower dietary intake. In an intervention study (McMurchie et al., 1984b) 16 subjects had their dietary polyunsaturated to saturated fat ratio increased from a mean of 0.27 to 1.06, principally by the replacement of saturated and monounsaturated fatty acids with linoleic acid. After 6 weeks, the mean plasma linoleic acid had increased by 9.0 percent of total plasma fatty acids, whereas linoleic acid in cheek cells had increased by 36.4 percent of total cheek cell fatty acids, however, there was no significant change among saturated fatty acids. These results are intriguing, but partially paradoxical, and more data are needed to define the utility of cheek cell fatty acid levels as measures of dietary intake.

Trans Fatty Acids

Concern has been expressed about the health effects of *trans* fatty acids in the modern diet. The only natural sources of *trans* fatty acids are foods containing ruminant fats (e.g., dairy products, beef fat). Industrial hydrogenation of vegetable oils (e.g., margarine) alters fatty acids from the *cis* isomer to the *trans* geometry (Ohlrogge et al., 1981). In a case-control study conducted in the United Kingdom, cases of myocardial infarction were observed to have higher values of *trans* unsaturated fatty acids in adipose tissue than controls (Thomas et al., 1983), although this analysis did not control for established risk factors for myocardial infarction.

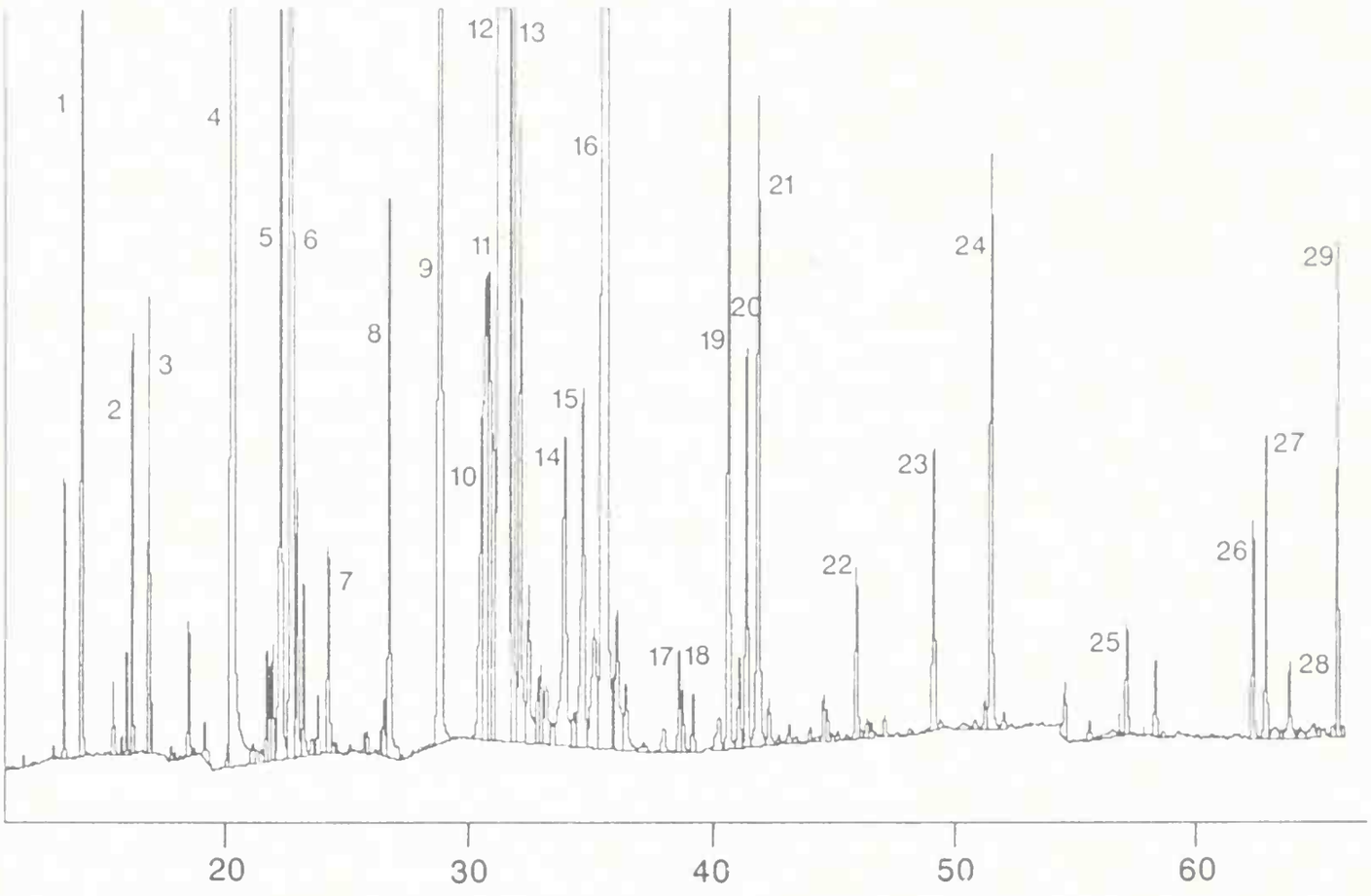
The average *trans* fatty acid content of the U.S. food supply is approximately 5 to 6 percent of total fat intake, which is similar to the *trans* fatty acid content of human adipose tissue (Senti, 1985). Similar agreement has been found in the United Kingdom, where the average *trans* intake of 5.4 percent of fat estimated in the average diet, was compared with an average of 5.2 percent of adipose fat obtained from 95 subjects (Thomas et al., 1981). No studies relating the dietary *trans* fatty acid intake of individuals to their *trans* fatty acid levels in adipose appear to have been undertaken.

Laboratory Techniques

If cholesteryl esters or phospholipids are used for fatty acid analysis, it is probably not necessary to collect fasting bloods. Subjects, however, should be fasting if the triglyceride fraction or free plasma levels are to be analyzed. The techniques of subcutaneous fat aspiration are well-described by Hirsch and colleagues (1960). Subsequent modifications of this technique eliminating the use of saline injection, and possibly local anesthetic, have made it even more applicable to large epidemiologic studies (Beynen and Katan, 1985a; Handelman, 1988). Use of glass apparatus and rigorous cleaning with organic solvents is most important. Separation of the triglyceride, cholesteryl esters, phospholipid, and free fatty acid fractions requires preparative thin layer chromatography. Gas-liquid chromatography (GLC) or high-performance liquid chromatography (HPLC) methods may be used to separate and measure the fatty acids within each fraction. Fatty acids are identified by their known chromatographic characteristics and by comparison with known standards, and they are quantitated by calculating the area under each peak. The sum of these individual fatty acids accounts for as much as 98 percent of the total peak area of fatty acids on the chromatogram. A typical high resolution chromatogram is displayed in Figure 9–7.

Stability

Beynen and Katan (1985a) observed no significant alterations in the fatty acid composition of abdominal adipose tissue stored at 4°C, −20°C, and −80°C for up to 18 months.



(1) 14:0 (Myristic Acid)
(2) 14:1ω5c
(3) 15:0
(4) 16:0 (Palmitic Acid)
(5) 16:1ω7t
(6) 16:1ω7c
(7) 17:0 (Margaric Acid)
(8) 17:1ω7c?
(9) 18:0 (Stearic Acid)
(10) 18:1ω9t
(11) 18:1ω7t
(12) 18:1ω9c (Oleic Acid)
(13) 18:1ω7c
(14) 18:2ω6tt
(15) 18:2ω6tc
(16) 18:2ω6cc (Linolenic Acid)
(17) 18:3ω6 (Gamma Linolenic Acid)
(18) 20:0
(19) 18:3ω3 (Linolenic Acid)
(20) 20:1ω9t?
(21) 20:1ω9c?
(22) 20:2ω6
(23) 20:3ω6 (Dihomogamma Linolenic Acid)
(24) 20:4ω6 (Arachidonic Acid)
(25) 20:5ω3 (Eicosapenataenoic Acid)
(26) 22:4ω6
(27) 24:1ω9
(28) 22:5ω3?
(29) 22:6ω3 (Docosahexaenoic Acid)

Figure 9–7. A high resolution gas-liquid chromatography chromatogram of a fatty acid profile.

PROTEIN

No single biochemical measure is adequate for the assessment of protein intake. In NHANES I, protein intake was a significant predictor of both serum total protein and serum albumin in multiple regression analyses. Even after adjustment for constitutional and sociodemographic factors, however, protein intake explained less than 10 percent of the variance in these biochemical measures (Kerr et al., 1982). In populations where protein intake is low or marginal, serum levels may be more useful. The other common measure of protein intake is urinary nitrogen excretion. Most of the nitrogen derived from amino acids produced by protein catabolism is excreted in urine. When dietary protein is reduced, these amino acids are used more efficiently, and thus less nitrogen is excreted (Young and Pellett, 1987). Excess ingested protein increases nitrogen excretion. In a study of eight healthy subjects, Bingham and Cummings (1985)

measured daily protein intake and urinary nitrogen excretion for 28 consecutive days. The mean within-person CV% in protein intake was 21 percent, whereas the mean within-person CV% in urinary nitrogen was substantially less at 13 percent. The high day-to-day variation in protein intake, reflected in the smaller, but still substantial variation in 24-hour urinary nitrogen, means that a single 24-hour urinary nitrogen will be a poor marker of long-term intake. Nitrogen excretion increases with body size and exercise, and decreased caloric intake results in a less efficient utilization of amino acids and higher urinary nitrogen excretion (Alpers et al., 1983). These nondietary influences limit the interpretability of urinary nitrogen as a measure of dietary protein.

Urinary measurement of 3-methylhistidine has been proposed as a specific indicator of meat consumption. The correlation between 24-hour urinary 3-methylhistidine and a food-frequency questionnaire meat score among 24 Canadian men was remarkably high ($r = 0.77$) (McKeown-Eyssen et al., 1986); this relationship needs to be examined further. The requirement for a 24-hour urine sample limits the usefulness of these markers for large epidemiologic studies.

FIBER

Although a decreased dietary fiber intake has been linked with a wide range of chronic diseases, few studies have been conducted using stool fiber content. Dietary intake of fiber includes crude fiber that is excreted unchanged, and soluble forms of fiber. Neutral detergent fiber (NDF) includes insoluble hemicellulose, cellulose, and lignin. Daily fecal weight, but not stool frequency, increases with increasing NDF consumption in isocaloric diets (Kelsay and Clark, 1984; Tucker et al., 1981). In a study in which total fiber intake was estimated from 2-day food records, significant correlations were found between fiber intake and fecal hemicellulose ($r = 0.54$) and percentage water ($r = 0.37$), and a nonsignificant positive correlation was observed with total fecal fiber ($r = 0.27$) (McKeown-Eyssen et al., 1986). Similar correlations were found with fiber intake estimated from a food-frequency questionnaire a year before the stool specimens were collected. These data suggest that total fiber, and its insoluble components measured in stool may be good indices of fiber intake. The difficulty of obtaining stool specimens from subjects in large epidemiologic studies severely limits the utility of this form of biochemical assessment.

SUMMARY

Biochemical measurements of nutrient levels in blood or other tissues can provide a useful assessment of the intake of certain nutrients. This approach is especially important for nutrients that are measured poorly by other methods due to large within-food variation in nutrient content.

The selection of a method for estimating intake of a particular nutrient must be made on a nutrient-by-nutrient basis. For several important nutrients, no feasible biochemical indicator of intake is available. For others, within-person

variation in level of the indicator, or the existence of other determinants, make correlations with long-term intake weak. For some nutrients, a biochemical indicator of intake exists which is sufficiently valid to be the method of choice. Careful attention to specimen collection, storage, and analysis, is vital to avoid misclassification or bias. As analytic methods improve and more biochemical indicators are validated as measures of dietary intake, their use in nutritional epidemiology is likely to expand.

REFERENCES

Aalberts, J. S., P. L. Weegels, L. van der Heijden, M. H. Borst, J. Burema, J. Hautvast, and T. Kouwenhoven (1988). Calcium supplementation: Effect on blood pressure and urinary mineral excretion in normotensive male lactoovovegetarians and omnivores. *Am. J. Clin. Nutr. 48,* 131–138.

Alpers, D. H., R. E. Clouse, and W. F. Stenson (1983). *Manual of nutritional therapeutics.* Boston, Mass.: Little, Brown and Co., p. 101.

Amedee-Manesme, O., M. S. Mourey, A. Hanck, and J. Therasse (1987a). Vitamin A relative dose response test: Validation by intravenous injection in children with liver disease. *Am. J. Clin. Nutr. 46,* 286–289.

Amedee-Manesme, O., R. Luzeau, C. Carlier, and A. Ellrodt (1987b). Simple impression cytology method for detecting vitamin A deficiency (letter). *Lancet i,* 1263.

Andon, M. B. and R. D. Reynolds (1987). A comparison of plasma pyridoxal 5′-phosphate concentrations in capillary (finger prick) and venous blood. *Am. J. Clin. Nutr. 45,* 1461–1465.

Arntzenius, A. C., D. Kromhout, J. D. Barth, J. H. Reiber, A. V. Bruschke, B. Buis, C. M. van Gent, N. Kempen-Voogd, S. Strikwerda, and E. van der Velde (1985). Diet, lipoproteins, and the progression of coronary atherosclerosis. *N. Engl. J. Med. 312,* 805–811.

Arroyave, G., C. O. Chichester, H. Flores, J. Glover, L. A. Mejia, J. A. Olson, K. L. Simpson, and B. A. Underwood (1982). *Report of IVACG.* Washington, D.C.: Nutrition Foundation.

Baker, H., O. Frank, and S. H. Hutner (1980). Vitamin analyses in medicine. In R. S. Goodhart, and M. E. Shils, eds.: *Modern Nutrition in Health and Disease.* Philadelphia, Pa.: Lea and Febiger, pp. 611–640.

Bamji, M. S. and F. Ahmed (1978). Effect of oral contraceptive steroids on vitamin status of women and female rats. *World Rev. Nutr. Diet 31,* 135–140.

Bankson, D. D., R. M. Russell, and J. A. Sadowski (1986). Determination of retinyl esters and retinol in serum or plasma by normal-phase liquid chromatography: Method and applications. *Clin. Chem. 32,* 35–40.

Basu, T. K., D. R. Patel, and D. C. Williams (1974). A simplified microassay of transketolase in human blood. *Int. J. Vitam. Nutr. Res. 44,* 319–326.

Basu, T. K. and C. J. Schorah (1982). Vitamin C reserves and requirements in health and disease. In T. K. Basu and C. J. Schorah, eds.: *Vitamin C in Health and Disease.* Westport, Conn.: AVI Publishing, pp. 61–92.

Bates, C. J., A. E. Black, D. R. Phillips, A. J. Wright, and D. A. Southgate (1982). The discrepancy between normal folate intakes and the folate RDA. *Hum. Nutr. Appl. Nutr. 36,* 422–429.

Beaton, G. H. and A. Chery (1986). Evaluation of methods used to assess dietary intake: Simulation analyses. *Can. J. Physiol. Pharmacol. 64,* 772–780.

Behrens, W. A. and R. Madere (1986). Alpha- and gamma-tocopherol concentrations in human serum. *J. Am. Coll. Nutr. 5*, 91–96.

Belsey, R. E., H. F. DeLuca, and J. T. Potts Jr. (1974). A rapid assay for 25-OH-vitamin D_3 without preparative chromatography. *J. Clin. Endocrinol. Metab. 38*, 1046–1051.

Bettendorff, L., C. Grandfils, C. de Rycker, and E. Schoeffeniels (1986). Determination of thiamine and its phosphate esters in human blood serum at femtomole levels. *J. Chromatog. 382*, 297–302.

Beynen, A. C. and M. B. Katan (1985a). Rapid sampling and long-term storage of subcutaneous adipose-tissue biopsies for determination of fatty acid composition. *Am. J. Clin. Nutr. 42*, 317–322.

Beynen, A. C. and M. B. Katan (1985b). Why do polyunsaturated fatty acids lower serum cholesterol? *Am. J. Clin. Nutr. 42*, 560–563.

Beynen, A. C., R. J. Hermus, and J. G. Hautvast (1980). A mathematical relationship between the fatty acid composition of the diet and that of the adipose tissue in man. *Am. J. Clin. Nutr. 33*, 81–85.

Bhagavan, H. N. (1985). Interaction between vitamin B-6 and drugs. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease*. N.Y.: Alan R. Liss, pp. 401–415.

Bieri, J. G., T. J. Tolliver, and G. L. Catignani (1979). Simultaneous determination of α-tocopherol and retinol in plasma or red cells by high pressure liquid chromatography. *Am. J. Clin. Nutr. 32*, 2143–2149.

Bieri, J. G., E. D. Brown, and J. C. Smith (1985). Determination of individual carotenoids in human plasma by high performance liquid chromatography. *J. Liquid Chromatogr. 8*, 473–484.

Biesalski, H., H. Greiff, K. Brodda, G. Hafner, and K. H. Bassler (1986). Rapid determination of vitamin A (retinol) and vitamin E (α-tocopherol) in human serum by isocratic adsorption HPLC. *Int. J. Vit. Nutr. Res. 56*, 319–327.

Bingham, S. A., J. H. Cummings (1985). Urine nitrogen as an independent validatory measure of dietary intake: A study of nitrogen balance in individuals consuming their normal diet. *Am. J. Clin. Nutr. 42*, 1276–1289.

Black, M. R., D. M. Medeiros, E. Brunett, and R. Welke (1988). Zinc supplements and serum lipids in young adult white males. *Am. J. Clin. Nutr. 47*, 970–975.

Blanchard, R. A., B. C. Furie, M. Jorgensen, S. F. Kruger, and B. Furie (1981). Acquired vitamin K-dependent carboxylation deficiency in liver disease. *N. Engl. J. Med. 305*, 242–248.

Boberg, M., L. B. Croon, I. B. Gustafsson, and B. Vessby (1985). Platelet fatty acid composition in relation to fatty acid composition in plasma and to serum lipoprotein lipids in healthy subjects with special reference to the linoleic acid pathway. *Clin. Science 68*, 581–587.

Bonanome, A. and S. M. Grundy (1988). Effect of dietary stearic acid on plasma cholesterol and lipoprotein levels. *N. Engl. J. Med. 318*, 1244–1248.

Brook M. and J. J. Grimshaw (1968). Vitamin C concentration of plasma and leukocytes as related to smoking habit, age, and sex of humans. *Am. J. Clin. Nutr. 21*, 1254–1258.

Bunker, V. W., L. J. Hinks, M. S. Lawson, and B. E. Clayton (1984). Assessment of zinc and copper status of healthy elderly people using metabolic balance studies and measurement of leucocyte concentrations. *Am. J. Clin. Nutr. 40*, 1096–1102.

Caggiula, A. W., R. R. Wing, M. P. Nowalk, N. C. Milas, S. Lee, and H. Langford (1985). The measurement of sodium and potassium intake. *Am. J. Clin. Nutr. 42*, 391–398.

Cameron, E. and L. Pauling (1978). Supplemental ascorbate in the supportive treatment

of cancer: prolongation of survival lines in terminal human cancer. *Proc. Natl. Acad. Sci. USA. 75,* 4538–4542.

Cameron, E., L. Pauling, and B. Leibovitz (1979). Ascorbic acid and cancer: A review. *Cancer Res. 39,* 663–681.

Camp, V. M., J. Chipponi, and B. A. Faraj (1983). Radioenzymatic assay for direct measurement of plasma pyridoxal-5′-phosphate. *Clin. Chem. 29,* 642–644.

Campos, F. A., H. Flores, and B. A. Underwood (1987). Effect of an infection on vitamin A status of children as measured by the relative dose response (RDR). *Am. J. Clin. Nutr. 46,* 91–94.

Castenmiller, J. J., R. P. Mensink, L. van der Heijden, T. Kouwenhoven, J. G. Hautvast, P. W. de Leeuw, and G. Schaafsma (1985). The effect of dietary sodium on urinary calcium and potassium excretion in normotensive men with different calcium intakes. *Am. J. Clin. Nutr. 41,* 52–60.

Cavill, I., A. Jacobs, and M. Worwood (1986). Diagnostic methods for iron status. *Ann. Clin. Biochem. 23,* 168–171.

Chen, X., G. Yang, J. Chen, X. Chen, Z. Wen, and K. Ge (1980). Studies on the relations of selenium and Keshan Disease. *Biol. Trace Element Res. 2,* 91–107.

Clark, L. C. (1985). The epidemiology of selenium and cancer. *Fed. Proc. 74,* 2584–2589.

Committee on Diet, Nutrition and Cancer, Assembly of Life Sciences, National Research Council (1982). In *Diet, Nutrition and Cancer.* Washington, D.C.: National Academy Press.

Comstock, G. W., M. S. Menkes, S. E. Schober, J.P. Vuilleumier, and K. J. Helsing (1988). Serum levels of retinol, beta-carotene, and alpha-tocopherol in older adults. *Am. J. Epidemiol. 127,* 114–123.

Cook, J. D. and B. S. Skikne (1982). Serum ferritin: A possible model for the assessment of nutrient stores. *Am. J. Clin. Nutr. 35,* 1180–1185.

Cooper, G. R. (1979). High density lipoprotein reference materials. In, K. Lippel, ed.: *Report of the High Density Lipoprotein Methodology Workshop.* DHEW Publ. No. 79-1661, Bethesda, Md.: U.S. Dept. of Health, Education and Welfare, Public Health Service, N.I.H. pp. 178–188.

Cousins, R. J. (1985). Absorption, transport, and hepatic metabolism of copper and zinc: Special reference to metallothionein and ceruloplasmin. *Physiol. Rev. 65,* 238–309.

Dayton, S., S. Hashimoto, W. Dixon, and M. L. Pearce (1966). Composition of lipids in human serum and adipose tissue during prolonged feeding of a diet high in unsaturated fat. *J. Lipid. Res. 7,* 103–111.

DeLuca, H. F. (1979). The vitamin D system in the regulation of calcium and phosphorus metabolism. *Nutr. Rev. 37,* 161–193.

Delves, H. T. (1976). The microdetermination of copper in plasma protein fractions. *Clin. Chem. Acta 71,* 495–500.

Dickson, I. (1987). New approaches to vitamin D. *Nature 325,* 18.

Dougherty, R. M., C. Galli, A. Ferro-Luzzi, and J. M. Iacono (1987). Lipid and phospholipid fatty acid composition of plasma, red blood cells, and platelets and how they are affected by dietary lipids: a study of normal subjects from Italy, Finland, and the U.S.A. *Am. J. Clin. Nutr. 45,* 443–455.

Driskell, J. A. (1978). Vitamin B-6 status of the elderly. In *National Research Council, Food and Nutrition Board. Human Vitamin B-6 Requirements.* Washington, D.C.: National Academy of Sciences, pp. 252–256.

Driskell, W. J., M. M. Bashor, and J. W. Neese (1985). Loss of vitamin A in long-term stored, frozen sera. *Clinica Chimica. Acta 147,* 25–30.

Driskell, J. A., A. J. Clark, T. L. Bazzarre, L. F. Chopin, H. McCoy, M. A. Kenney, and

S. W. Mack (1985). Vitamin B-6 status of southern adolescent girls. *J. Am. Diet. Assoc. 85,* 46–49.

Dyerberg, J., H. O. Bang, and N. Hjorne (1975). Fatty acid composition of the plasma lipids in Greenland Eskimos. *Am. J. Clin. Nutr. 28,* 958–966.

Dyerberg, J. and H. O. Bang (1979). Haemostatic function and platelet polyunsaturated fatty acids in Eskimos. *Lancet ii,* 433–435.

Farleigh, C. A., R. Shepherd, and D. G. Land (1985). Measurement of sodium intake and its relationship to blood pressure and salivary sodium concentration. *Nutr. Res. 5,* 815–826.

Farquhar, J. W. and E. H. Ahrens, Jr. (1963). Effects of dietary fats on human erythrocyte fatty acid patterns. *J. Clin. Invest. 42,* 675–685.

Field, C. J., A. Angel, and M. T. Clandenin (1985). Relationship of diet to the fatty acid composition of human adipose tissue structural and stored lipids. *Am. J. Clin. Nutr. 42,* 1206–1220.

Fischer, P. W., A. Giroux, and M. R. L'Abbe (1984). Effect of zinc supplementation on copper status in adult men. *Am. J. Clin. Nutr. 40,* 743–746.

Fordyce, M. K., G. Christakis, A. Kafatos, R. Duncan, and J. Cassady (1983). Adipose tissue fatty acid composition of adolescents in a U.S.–Greece cross-cultural study of coronary heart disease risk factors. *J. Chronic Dis. 36,* 481–486.

Frassinelli-Gunderson, E. P., S. Margen, and J. R. Brown (1985). Iron stores in users of oral contraceptive agents. *Am. J. Clin. Nutr. 41,* 703–712.

Garland, C., R. B. Shekelle, E. Barrett-Connor, M. H. Criqui, A. H. Rossof, and O. Paul (1985). Dietary vitamin D and calcium and risk of colorectal cancer: A 19-year prospective study in men. *Lancet (i),* 307–309.

Gleuck, C. J., D. J. Gordon, J. J. Nelson, C. E. Davis, and H. A. Tyroler (1986). Dietary and other correlates of changes in total and low density lipoprotein cholesterol in hypercholesterolemic men. The lipid research clinics coronary primary prevention trial. *Am. J. Clin. Nutr. 44,* 489–500.

Gordon, T., W. P. Castelli, M. C. Hjortland, W. B. Kannel, and T. R. Dawber (1977). High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *Am. J. Med. 62,* 707–714.

Griffiths, N. M., R. D. Stewart, and M. F. Robinson (1976). The metabolism of [^{75}Se] selenomethionine in four women. *Br. J. Nutr. 35,* 373–382.

Grundy, S. M. and G. L. Vega (1988). Plasma cholesterol responsiveness to saturated fatty acids. *Am. J. Clin. Nutr. 47,* 822–824.

Gurr, M. I., R. T. Jung, M. P. Robinson, and W. P. James (1982). Adipose tissue cellularity in man: The relationship between fat cell size and number, the mass and distribution of body fat and the history of weight gain and loss. *Int. J. Obesity 6,* 419–436.

Haddad, J. G. Jr. and T. J. Hahn (1973). Natural and synthetic sources of circulating 25-hydroxyvitamin D in man. *Nature 244,* 515–516.

Hainline, A. Jr., C. L. Winn, G. R. Cooper, D. T. Miller, and D. D. Bayse (1978). The CDC cooperative cholesterol and triglyceride standardization program—Twenty years experience. *Clin. Chem. 24,* 1020.

Hallberg, L., L. Rossander (1984). Improvement of iron nutrition in developing countries: comparison of adding meat, soy protein, ascorbic acid, citric acid, and ferrous sulphate on iron absorption from a simple Latin American-type of meal. *Am. J. Clin. Nutr. 39,* 577–583.

Hambidge, K. M. (1973). Increase in hair copper concentration with increasing distance from the scalp. *Am. J. Clin. Nutr. 26,* 1212–1215.

Handelman, G. J., W. L. Epstein, L. J. Machlin, F.J.G.M. van Kuijk, and E. A. Dratz (1988). Biopsy method for human adipose with vitamin E and lipid measurements. *Lipids 23*, 598–604.

Haussler, M. R. and T. A. McCain (1977). Basic and clinical concepts related to vitamin D metabolism and action. *N. Engl. J. Med. 297*, 974–983.

Hegsted, D. M., R. B. McGandy, M. L. Myers, and F. J. Stare (1965). Quantitative effects of dietary fat on serum cholesterol in man. *Am. J. Clin. Nutr. 17*, 281–295.

Hegsted, D. M. and R. J. Nicolosi (1987). Individual variation in serum cholesterol levels. *Proc. Natl. Acad. Sci. USA 84*, 6259–6261.

Herbert, V. (1987). Recommended dietary intakes (RDI) of folate in humans. *Am. J. Clin. Nutr. 45*, 661–670.

Herbeth, B., L. Didelot-Barthelemy, A. Lemoine, and C. Le Devehat (1988). Plasma fat-soluble vitamins and alcohol consumption (letter). *Am. J. Clin. Nutr. 47*, 343–344.

Hirsch, J., J. W. Farquhar, E. H. Ahrens, M. L. Peterson, and W. Stoffel (1960). Studies of adipose tissue in man; A microtechnic for sampling and analysis. *Am. J. Clin. Nutr. 8*, 499–511.

Holbrook, J. T., K. Y. Patterson, J. E. Bodner, L. W. Douglas, C. Veillon, J. L. Kelsay, W. Mertz, and J. C. Smith Jr. (1984). Sodium and potassium intake and balance in adults consuming self-selected diets. *Am. J. Clin. Nutr. 40*, 786–793.

Holman, R. T. (1986). Control of polyunsaturated acids in tissue lipids. *J. Am. Coll. Nutr. 5*, 183–211.

Horwitt, M. K. (1987). Human requirements for riboflavin: Reply to letter by Bates. *Am. J. Clin. Nutr. 46*, 123.

Howard, M. P., M. A. Andon, and R. D. Reynolds (1984). Long-term stability of pyridoxal phosphate in frozen human plasma. *Fed. Proc. 43*, 486.

Hughes, M. R., D. J. Baylink, P. G. Jones, and M. R. Haussler (1976). Radioligand receptor assay for 25-hydroxyvitamin D-2/D-3 and 1 α, 25-dihydroxyvitamin D-2/D-3. *J. Clin. Invest. 58*, 61–70.

Hulley, S. B., R. Cohen, and G. Widdowson (1977). Plasma high density lipoprotein cholesterol level: Influence of risk factor intervention. *J. Am. Med. Assoc. 238*, 2269–2271.

Hunt, I. F., N. J. Murphy, A. E. Cleaver, B. Faraji, M. E. Swendseid, A. H. Coulson, V. A. Clark, N. Laine, C. A. Davis, and J. C. Smith Jr. (1983). Zinc supplementation during pregnancy: Zinc concentration of serum and hair from low-income women of Mexican descent. *Am. J. Clin. Nutr. 37*, 572–582.

Hunter, D. J., C. G. Chute, E. Kushner, G. A. Colditz, M. J. Stampfer, F. E. Speizer, J. S. Morris, and W. C. Willett (1987). Predictors of selenium concentration in nail tissue. *Am. J. Epidemiol. 126*, 743.

Iber, F. L., J. P. Blass, M. Brin, and C. M. Leevy (1982). Thiamin in the elderly—Relation to alcoholism and to neurological degenerative disease. *Am. J. Clin. Nutr. 36*, 1067–1082.

Interdepartmental Committee on Nutrition for National Defense (1963). *Manual for Nutritional Surveys*, 2nd ed. Bethesda, Md.: National Institutes of Health.

International Collaborative Study Group (1986). Metabolic epidemiology of plasma cholesterol: mechanisms of variation of plasma cholesterol within populations and between populations. *Lancet ii*, 991–996.

Irwin, M. I. and B. K. Hutchins (1976). A conspectus of research on vitamin C requirements of man. *J. Nutr. 106*, 821–879.

Iyengar, V. (1987). Dietary intake studies of nutrients and selected toxic elements in human studies: Analytical approaches. *Clin. Nutr. 6*, 105–117.

Jacob, R. A., J. H. Skala, and S. J. Omaye (1987). Biochemical indices of human vitamin C status. *Am. J. Clin. Nutr. 46,* 818–826.

Jacobs, D. R. Jr. and E. Barrett-Connor (1982). Retest reliability of plasma cholesterol and triglyceride. The Lipid Research Clinics Prevalence Study. *Am. J. Epidemiol. 116,* 878–885.

Joffres, M. R., D. M. Reed, and K. Yano (1987). Relationship of magnesium intake and other dietary factors to blood pressure: The Honolulu heart study. *Am. J. Clin. Nutr. 45,* 469–475.

Kalman, D. A., G. E. Goodman, G. S. Osmenn, G. Bellamy, and B. Rollins (1987). Micronutrient assay for cancer prevention clinical trials: Serum retinol, retinyl palmitate, alpha-carotene, and beta-carotene with the use of high-performance liquid chromatography. *J. N. C. I. 79,* 975–982.

Katan, M. B., A. C. Beynen, J. H. de Vries, and A. Nobels (1986). Existence of consistent hypo- and hyperresponders to dietary cholesterol in man. *Am. J. Epidemiol. 123,* 221–234.

Kelsay, J. L. and W. M. Clark (1984). Fiber intakes, stool frequency, and stool weights of subjects consuming self-selected diets. *Am. J. Clin. Nutr. 40,* 1357–1360.

Kerr, G. R., E. S. Lee, M. K. Lam, R. J. Lorimor, E. Randall, R. N. Forthofer, M. A. Davis, and S. M. Magnetti (1982). Relationship between dietary and biochemical measures of nutritional status in HANES I data. *Am. J. Clin. Nutr. 35,* 294–307.

Keshan Disease Research Group of the Chinese Academy of Medical Sciences (1979). Epidemiologic studies on the etiologic relationship of selenium and Keshan disease. *Chin. Med. J. 92,* 477–482.

Keys, A. (1984). Serum-cholesterol response to dietary cholesterol. *Am. J. Clin. Nutr. 40,* 351–359.

Keys, A., J. T. Anderson, and G. Grande (1965). Serum cholesterol response to changes in the diet. *Metabolism 14,* 747–787.

Khaw, K. T. and E. Barrett-Connor (1987). Dietary potassium and stroke-associated mortality. A 12-year prospective population study. *N. Engl. J. Med. 316,* 235–240.

Kjeldsen, S. E., I. Eide, P. Leren, and O. P. Foss (1983). Effects of posture on serum cholesterol fractions, cholesterol ratio and triglycerides. *Scand. J. Clin. Lab. Invest. 43,* 119–121.

Kjosen, B. and S. H. Seim (1977). The transketolase assay of thiamin in some diseases. *Am. J. Clin. Nutr. 30,* 1591–1596.

Klasing, S. A. and S. M. Pilch, eds. (1985). *Suggested Measures of Nutritional Status and Health Conditions for the Third National Health and Nutrition Examination Survey.* Bethesda, Md.: Federation of American Societies for Experimental Biology.

Klevay, L. M. (1983). Copper and ischemic heart disease. *Biol. Trace Element Res. 5,* 245–255.

Klevay, L. M., S. J. Reck, R. A. Jacob, G. M. Logan, Jr., J. M. Munoz, and H. H. Sandstead (1980). The human requirement for copper. I. Healthy men fed conventional American diets. *Am. J. Clin. Nutr. 33,* 45–50.

Knekt, P., A. Aromaa, J. Maatela, R.K. Aaran, T. Nikkari, M. Hakama, T. Hakulinen, R. Peto, E. Saxen, and L. Teppo (1988). Serum vitamin E and risk of cancer among Finnish men during a 10-year follow-up. *Am. J. Epidemiol. 127,* 28–41.

Korpela, H., J. Kumpulainen, P. V. Luoma, A. J. Arranto, E. A. Sotaniemi (1985). Decreased serum selenium in alcoholics as related to liver structure and function. *Am. J. Clin. Nutr. 42,* 147–151.

Kushi, L. H., R. A. Lew, F. J. Stare, C. R. Ellison, M. el Lozy, G. Bourke, L. Daly, I. Graham, N. Hickey, R. Mulcahy, et al. (1985). Diet and 20-year mortality from cor-

onary heart disease. The Ireland-Boston Diet-Heart Study. *N. Engl. J. Med. 312,* 811–818.

Kushi, L. H., K. W. Samonds, J. M. Lacey, P. T. Brown, J. G. Bergan, and F. M. Sacks (1988). The association of dietary fat with serum cholesterol in vegetarians: the effect of dietary assessment on the correlation coefficient *Am. J. Epidemiol. 128,* 1054–1064.

Lehmann, J., D. D. Rao, J. J. Canary, and J. T. Judd (1988). Vitamin E and relationships among tocopherols in human plasma, platelets, lymphocytes, and red blood cells. *Am. J. Clin. Nutr. 47,* 470–474.

Leklem, J. E. (1985). Physical activity and vitamin B-6 metabolism in men and women: interrelationship with fuel needs. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* N.Y.: Alan R. Liss, pp. 221–241.

Lemoyne, M., A. Van Gossum, R. Kurian, M. Ostro, J. Axler, and K. N. Jeejeebhoy. (1987). Breath pentane analysis as an index of lipid peroxidation: A functional test of vitamin E status. *Am. J. Clin. Nutr. 46,* 267–272.

Levander, O. A. (1985). Considerations on the assessment of selenium status. *Fed. Proc. 44,* 2579–2583.

Lind, J. (1753). *A Treatise on the Scurvy.* London: A. Millar. (Republished Edinburgh, Edinburgh University Press, 1953).

Linscheer, W. G. and A. J. Vergroesen (1988). Lipids. In M. E . Shils, and V. R. Young, eds.: *Modern Nutrition in Health and Disease* Philadelphia, Pa.: Lea and Febiger, pp. 72–107.

Lips, P., F. C. van Ginkel, M. J. Jongen, F. Rubertus, W. J. van der Vijgh, and J. C. Netelenbos (1987). Determinants of vitamin D status in patients with hip fracture and in elderly control subjects. *Am. J. Clin. Nutr. 46,* 1005–1010.

Liu, K., R. Cooper, J., McKeever, P. McKeever, R. Byington, I. Soltero, R. Stamler, F. Gosch, E. Stevens, and J. Stamler (1979a). Assessment of the association between habitual salt intake and high blood pressure: Methodological problems. *Am. J. Epidemiol. 110,* 219–226.

Liu, K., R. Cooper, I. Soltero, and J. Stamler (1979b). Variability in 24-hour urine sodium excretion in children. *Hypertension 1,* 631–636.

Lloyd, B., R. S. Lloyd, B. E. Clayton (1983). Effects of smoking, alcohol and other factors on the selenium status of a healthy population. *J. Epidemiol. Com. Health. 37,* 213–217.

Loerch, J. D., B. A. Underwood, and K. C. Lewis (1979). Response of plasma levels of vitamin A to a dose of vitamin A as an indicator of hepatic vitamin A reserves in rats. *J. Nutr. 109,* 778–786.

Loh, H. S. (1972). The relationship between dietary ascorbic acid intake and buffy coat and plasma ascorbic acid concentrations at different ages. *Internat. J. Vit. Nutr. Res. 42,* 80–85.

Longnecker, M. P., P. R. Taylor, O. A. Levander, S. M. Howe, C. Veillon, P. A. McAdam, K. Y. Patterson, J. M. Holden, M. J. Stampfer, J. S. Morris, and W. C. Willett (1987). Tissue Selenium (Se) levels and indices of Se exposure in a seleniferous area. *Fed. Proc. 46,* 587 (abs).

Luoma, H., A. Aromaa, S. Helminen, N. Murtomaa, L. Kiviluoto, S. Punsar, and P. Knekt (1983). Risk of myocardial infarction in Finnish men in relation to fluoride, magnesium and calcium concentration in drinking water. *Acta. Med. Scand. 213,* 171–176.

Luoma, P. V., H. Korpela, E. A. Sotaniemi, and J. Kumpulainen (1985). Serum selenium, glutathione peroxidase, lipids, and human liver microsomal enzyme activity. *Biol. Tr. El. Res. 8,* 113–121.

Lumeng, L., T. K. Li, and A. Lui (1985). Interorgan transport and metabolism of vitamin B-6. In R. D. Reynolds, J. F. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* N.Y.: Alan R. Liss, pp. 35–54.

MacMahon, B. and T. F. Pugh (1970). *Epidemiology: Principles and Methods,* 1st ed. Boston, Mass.: Little, Brown and Company.

Manku, M. S., D. F. Horrobin, Y. S. Huang, and N. Morse (1983). Fatty acids in plasma and red cell membranes in normal humans. *Lipids 18,* 906–908.

Manore, M. N., J. E. Leklem, and M. C. Walter (1987). Vitamin B-6 metabolism as affected by exercise in trained and untrained women fed diets differing in carbohydrate and vitamin B-6 content. *Am. J. Clin. Nutr. 46,* 995–1004.

Martin, G. M. (1964). Copper content of hair and nails of normal individuals and of patients with hepatolenticular degeneration. *Nature 202,* 903–904.

McAdam, P. A., V. C. Morris, and O. A. Levander (1984). Automated determination of glutathione peroxidase (GSU-Px) activity in tissue from rats of different selenium (Se) status (abstr.). *Fed. Proc. 43,* 867.

McCarron, D. A., C. D. Morris, H. J. Henry, and J. L. Stanton (1984). Blood pressure and nutrient intake in the United States. *Science 224,* 1392–1398.

McCormick, D. B. (1985). Vitamins. In N. W. Tietz, ed.: *Textbook of Clinical Chemistry.* Philadelphia, Pa.: W. B. Saunders.

McKenna, M. J., R. Freaney, A. Meade, and F. P. Muldowney (1985). Hypovitaminosis D and elevated serum alkaline phosphatase in elderly Irish people. *Am. J. Clin. Nutr. 41,* 101–109.

McKeown-Eyssen, G. E., K. S. Yeung, and E. Bright-See (1986). Assessment of past diet in epdemiologic studies. *Am. J. Epidemiol. 124,* 94–103.

McLeroy, V. J. and H. E. Schendel (1973). Influence of oral contraceptives on ascorbic acid concentrations in healthy, sexually mature women. *Am. J. Clin. Nutr. 26,* 191–196.

McMurchie, E. J., J. D. Potter, T. E. Rohan, B. S. Hetzel (1984a). Human cheek cells; a non-invasive method for determining tissue lipid profiles in dietary and nutritional studies. *Nutr. Rep. Int. 29,* 519–526.

McMurchie, E. J., B. M. Margetts, L. J. Beilin, K. D. Croft, R. Vandongen, and B. K. Armstrong (1984b). Dietary-induced changes in the fatty acid composition of human cheek cell phospholipids: Correlation with changes in the dietary polyunsaturated/saturated fat ratio. *Am. J. Clin. Nutr. 39,* 965–980.

McNamara, D. J., R. Kolb, T. S. Parker, H. Batwin, P. Samuel, C. D. Brown, and E. H. Ahrens Jr. (1987). Heterogeneity of cholesterol homeostasis in man. Response to changes in dietary fat quality and cholesterol quantity. *J. Clin. Invest. 79,* 1729–1739.

Mejia, L. A. and G. Arroyave (1983). Determination of vitamin A in blood. Some practical considerations on the time of collection of the specimens and the stability of the vitamin. *Am. J. Clin. Nutr. 37,* 147–151.

Miller, L. T. (1985). Oral contraceptives and vitamin B-6 metabolism. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* N.Y.: Alan R. Liss, pp. 243–255.

Milne, D. B., S. Gallagher, C. Stjern, L. M. Klevay, and H. H. Sandstead (1982). Superoxide dismutase activity as an index of copper nutriture in man (abstr.). *Fed. Proc. 41,* 785.

Mino, M. and M. Nagamato (1986). An evaluation of nutritional status of vitamin E in pregnant women with respect to red blood cell tocopherol level. *Int. J. Vit. Nutr. Res. 56,* 149–153.

Mjos, O. D., S. N. Rao, L. Bjoru, T. Henden, D. S. Thelle, O. H. Forde, and N. E. Miller

(1979). A longitudinal study of the biological variability of plasma lipoproteins in healthy young adults. *Atherosclerosis 34,* 75–81.

Morris, J. S., M. J. Stampfer, and W. Willett (1983). Dietary selenium in humans: Toenails as an indicator. *Biol. Trace Element Res. 5,* 529–537.

Naito, H. K. and J. A. David (1984). Laboratory considerations: Determination of cholesterol, triglyceride, phospholipid, and other lipids in blood and tissues. In. J. A. Storey, ed.: *Lipid Research Methodology.* N.Y.: Alan R. Liss, Inc. pp. 1–76.

National Diet Heart Study Research Group (1968). The National Diet Heart Study Final Report. Chapter XII. Serum Cholesterol Response. I:181–223. American heart Association Monograph No. 18. New York, N.Y.: American Heart Association.

Neve, J., F. Vertongen, and P. Capel (1988). Selenium supplementation in healthy Belgian adults: Response in platelet glutathione peroxidase activity and other blood indices. *Am. J. Clin. Nutr. 48,* 139–143.

Newton, H. M., M. Sheltawy, A. W. Hay, and B. Morgan (1985). The relations between vitamin D_2 and D_3 in the diet and plasma $25OHD_2$ and $25OHD_3$ in elderly women in Great Britain. *Am. J. Clin. Nutr. 41,* 760–764.

Nierenberg, D. W. (1984). Determination of serum and plasma concentrations of retinol using high-performance liquid chromatography. *J. Chromatogr. 311,* 239–248.

Norris, R. L., M. J. Thomas, and P. W. Craswell (1986). Assessment of a two-step high-performance liquid chromatographic assay using dual-wavelength ultraviolet monitoring for 25-hydroxyergocalciferol and 25-hydroxycholecalciferol in human serum or plasma. *J. Chromatogr. 381,* 53–61.

O'Dell, B. L. (1984) Bioavailability of trace elements. *Nutr. Rev. 42,* 301–308.

Ohlrogge, J. B., E. A. Emken, and R. M. Gulley (1981). Human tissue lipids: Occurrence of fatty acid isomers from dietary hydrogenated oils. *J. Lipid Res. 22,* 955–960.

Olson, J. A. (1984). Serum levels of vitamin A and carotenoids as reflectors of nutritional status. *J. N. C. I. 73,* 1439–1444.

Olson, J. A. (1987). Recommended dietary intakes (RDI) of vitamin K in humans. *Am. J. Clin. Nutr. 45,* 687–692.

Olson, R. E. (1980). Vitamin K. In. R. S. Goodhart, M. G. Shils, eds.: *Modern Nutrition in Health and Disease.* Philadelphia, Pa.: Lea and Febiger, pp. 170–180.

Olson, J. A. and R. E. Hodges (1987). Recommended dietary intakes (RDI) of vitamin C in humans. *Am. J. Clin. Nutr. 45,* 693–703.

Omdahl, J. L. (1978). Interaction of the parathyroid and 1,25-dihydroxyvitamin D_3 in the control of renal 25-hydroxyvitamin D_3 metabolism. *J. Biol. Chem. 253,* 8474–8478.

Omdahl, J. L., P. J. Garry, L. A. Hunsaker, W. C. Hunt, and J. S. Goodwin. (1982). Nutritional status in a healthy elderly population: Vitamin D. *Am. J. Clin. Nutr. 36,* 1225–1233.

Pachla, L. A., D. L. Reynolds, and P. T. Kissinger (1985). Analytical methods for determining ascorbic acid in biological samples, food products, and pharmaceuticals. *J. Assoc. Off. Anal. Chem. 68,* 1–12.

Parfitt, A. M., J. C. Gallagher, R. P. Heaney, C. C. Johnston, R. Neer, and G. D. Whedon (1982). Vitamin D and bone health in the elderly. *Am. J. Clin. Nutr. 36,* 1014–1031.

Parker, R. S. (1988). Carotenoid and tocopherol composition of human adipose tissue. *Am. J. Clin. Nutr. 47,* 33–36.

Parr, R.M. (1985). Quality assurance of trace element analyses. *Nutr. Res. Suppl(1),* 5–11.

Patwardhan, V. N. (1969). Hypovitaminosis A and epidemiology of xerophthalmia. *Am. J. Clinc. Nutr. 22,* 1106–1118.

Pauling, L. (1971). The significance of the evidence about ascorbic acid and the common cold. *Proc. Natl. Acad. Sci. USA. 68,* 2678–2681.

Peng, Y. M., M. J. Xu, and D. S. Alberts (1987). Analysis and stability of retinol in plasma. *J. N. C. I. 78,* 95–99.

Peto, R. (1983). The marked differences between carotenoids and retinoids: Methodological implications for biochemical epidemiology. *Cancer Surveys 2,* 327–340.

Peto, R., R. Doll, J. D. Buckley, and M. B. Sporn (1981). Can dietary beta-carotene materially reduce human cancer rates? *Nature 290,* 201–208.

Phillips, G. B. and J. T. Dodge (1967). Composition of phospholipids and phospholipid fatty acids of human plasma. *J. Lipid Res. 8,* 676–681.

Pilch, S. M., ed. (1985). *Assessment of the Vitamin A Nutritional Status of the U. S. Population Based on Data Collected in the Health and Nutrition Examination Surveys.* Bethesda, Md.: Life Sciences Research Office, Federation of American Societies for Experimental Biology.

Pilch, S. M. and F. R. Senti (1984a). In *Assessment of the Iron Nutritional Status of the U.S. Population Based on Data Collected in the Second National Health and Nutrition Examination Survey, 1976–1980.* Bethesda, Md.: Life Sciences Research Office, Federation of American Societies for Experimental Biology, pp. 19–23.

Pilch, S. M. and F. R. Senti, eds.. (1984b). *Assessment of the Zinc Nutritional Status of the U.S. Population Based on Data Collected in the Second National Health and Nutrition Examination Survey, 1976–1980.* Bethesda, Md.: Life Sciences Research Office, Federation of American Societies for Experimental Biology.

Pilon, V. A., P. J. Howanitz, J. H. Howanitz, and N. Domres (1981). Day-to-day variation in serum ferritin concentration in healthy subjects. *Clin. Chem. 27,* 78–82.

Prasad, A. S., P. Rabbani, A. Abbasii, et al., (1978). Experimental zinc deficiency in humans. *Ann. Intern. Med. 89,* 483–490.

Pratt, W. B., J. L. Omdahl, and J.R.J. Sorenson (1985). Lack of effects of copper gluconate supplementation. *Am. J. Clin. Nutr. 42,* 681–682.

Rea, H. M., C. D. Thomson, D. R. Campbell, and M. F. Robinson (1979). Relation between erythrocyte selenium concentrations and glutathione peroxidase activities of New Zealand residents and visitors to New Zealand. *Br. J. Nutr. 42,* 201–208.

Reeves, V. B., E. J. Matausik, and J. L. Kelsay (1984). Variations in plasma fatty acid concentrations during a one-year self-selected dietary intake study. *Am. J. Clin. Nutr. 40,* 1345–1351.

Reiser, S., J. C. Smith Jr., W. Mertz, J. J. Holbrook, D. J. Scholfield, A. S. Powell, W. K. Canfield, and J. J. Canary (1985). Indices of copper status in humans consuming a typical American diet containing either fructose or starch. *Am. J. Clin. Nutr. 42,* 242–251.

Reynolds, R. D. and J. E. Leklem (1985). Implications on the role of vitamin B-6 in health and disease—A summary. In. R. D. Reynolds, J. F. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* N.Y.: Alan R. Liss, pp. 481–489.

Riboli, E., H. Ronnholm, and R. Saracci (1987). Biologic markers of diet. *Cancer Surveys 6,* 685–718.

Robinson, J. R., M. F. Robinson, O. A. Levander, and C. D. Thomson (1985). Urinary excretion of selenium by New Zealand and North American human subjects on differing intakes. *Am. J. Clin. Nutr. 41,* 1023–1031.

Robinson, M. F. and C. D. Thomson (1981). Selenium levels in humans vs environmental sources. In Spallholz, J. E., et al., eds.: *Selenium in Biology and Medicine.* Westport, Conn. AVI Publishing Company, Inc., pp. 283–302.

Russell, M. J., B. S. Thomas, and R. D. Bulbrook (1988). A prospective study of the relationship between serum vitamins A and E and risk of breast cancer. *Br. J. Cancer 57,* 213–215.

Russell, R. M., F. L. Iber, S. D. Kransinski, and P. Miller (1983). Protein–energy malnutrition and liver dysfunction limit the usefulness of the relative dose response (RDR) test for predicting vitamin A deficiency. *Hum. Nutr. Clin. Nutr. 37,* 361–371.

Russell, R. M., S. D. Krasinski, and B. Dawson-Hughes (1984). Indices of fat-soluble vitamin states. *Clin. Nutr. 3,* 161–168.

Russell-Briefel, R., M. W. Bates and L. H. Kuller (1985). The relationship of plasma carotenoids to health and biochemical factors in middle-aged men. *Am. J. Epidemiol. 122,* 741–749.

Sacks, F. M., D. Ornish, B. Rosner, S. McLanahan, W. P. Castelli, and E. M. Kass (1985). Plasma lipoprotein levels in vegetarians. The effect of ingestion of fats from dairy products. *J.A.M.A. 254,* 1337–1341.

Sacks, F. M., M. J. Stampfer, A. Munoz, K. McManus, M. Canessa, and E. H. Kass (1987). Effect of linoleic acid and oleic acids on blood pressure, blood viscosity, and erythrocyte cation transport. *J. Am. Coll. Nutr. 6,* 179–185.

Salbe, A. D. and O. A. Levander (1987). Hair and nails as indicators of selenium (Se) status in rats fed elevated dietary levels of Se as L-selenomethionine (Se Met) or sodium selenate (Na_2SeO_4). *Fed. Proc. 46,* 1153.

Samuel, P., D. J. McNamara, and J. Shapiro (1983). The role of diet in the etiology and treatment of atherosclerosis. *Ann. Rev. Med. 34,* 179–194.

Sandstead, N. N., L, K. Henrikson, J. L. Greger, A. S. Prasad, and R. A. Good (1982). Zinc nutriture in the elderly in relation to taste acuity, immune response, and wound healing. *Am. J. Clin. Nutr. 36,* 1046–1059.

Sauberlich, H. E. (1984). Newer laboratory methods for assessing nutriture of selected B-complex vitamins. *Annu. Rev. Nutr. 4,* 377–407.

Sauberlich, H. E., J. H. Skala, and R. P. Dowdy (1974). In *Laboratory Tests for the Assessment of Nutritional Status.* Cleveland, Ohio: CRC Press Inc., p. 7.

Schrauzer, G. N., D. A. White, and C. J. Schneider (1977). Cancer mortality correlation studies—III: Statistical associations with dietary selenium intakes. *Bioinorg. Chem. 7,* 23–31.

Schultz, T. D. and J. E. Leklem (1981). Urinary 4-pyridoxic acid, urinary vitamin B-6 and plasma pyridoxal phosphate as measures of vitamin B-6 status and dietary intake of adults. In J. E. Leklem, R. D. Reynolds, eds.: *Methods in Vitamin B-6 Nutrition: Analysis and Status Assessment.* New York, N.Y.: Plenum Press, pp. 297–320.

Schultz, T. D. and J. E. Leklem (1985). Supplementation and vitamin B-6 metabolism. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* New York, N.Y.: Alan R. Liss, pp. 419–427.

Senti, F. R. (1985). In *Health Aspects of Dietary* trans *Fatty Acids.* Bethesda, Md.: Life Sciences Research Office, Federation of American Societies for Experimental Biology.

Senti, F. R. and S. M. Pilch (1984). In F. R. Senti, and S. M. Pilch, eds.: *Assessment of the Folate Nutritional Status of the U.S. Population Based on Data Collected in the Second National Health and Nutritional Examination Survey, 1976–1980.* Bethesda, Md.: Life Sciences Research Office, Federation of American Societies for Experimental Biology.

Shekelle, R. B. and A. M. Shryock, O. Paul, M. Lepper, J. Stamler, S. Liu, and W. J. Raynor Jr. (1981). Diet, serum cholesterol, and death from coronary heart disease. *N. Engl. J. Med. 304,* 65–70.

Shearer, M. J., J. Oyeyi, S. Rahim, Y. Haroon, and P. Barkhan (1982). Endogenous vitamin K_1 in human plasma measured by high-performance liquid chromatography. *J. Br. Soc. Haematol. 50,* 690.

Sherwin, R. W., D. N. Wentworth, J. A. Cutler, S. B. Hulley, L. H. Kuller, and J. Stamler

(1987). Serum cholesterol levels and cancer mortality in 361,662 men screened for the Multiple Risk Factor Intervention Trial. *J.A.M.A. 257,* 943–948.

Silverman, D. I., G. J. Reis, F. M. Sacks, T. M. Boucher, M. E. Slipperly, and R. C. Pasternak (1988). Plasma phospholipid EPA levels: A new measure of dietary fish consumption. *Circulation Suppl. 78,* –227

Sinclair, H. and M. Gale (1987). Eicosopentaenoic acid in fat (letter). *Lancet i,* 1202.

Smith, F. R. and D. S. Goodman (1976). Vitamin A transport in human vitamin A toxicity. *N. Engl. J. Med.* 294, 805–808.

Smith, J. C., J. T. Holbrook, and D. E. Danford (1985). Analysis and evaluation of zinc and copper in human plasma and serum. *J. Am. Coll. Nutr. 4,* 627–638.

Smithells, R. W., N. C. Nevin, M. J. Seller, S. Sheppard, R. Harris, A. P. Read, D. W. Fielding, S. Walker, C. J. Schorah, and J. Wilde (1983). Further experience of vitamin supplementation for prevention of neural tube defect recurrences. *Lancet i,* 1027–1031.

Sokol, R. J., J. E Heubi, S. T. Iannaccone, K. E. Bove, W. F. Balistreri (1984). Vitamin E deficiency with normal serum vitamin E concentrations in children with chronic cholestasis. *N. Engl. J. Med. 310,* 1209–1212.

Solomons, N. W. (1985). Biochemical, metabolic, and clinical role of copper in human nutrition. *J. Am. Coll. Nutr. 4,* 83–105.

Solomons, N. W., T. J. Layden, I. H. Rosenberg, K. Vo-Khactu, and H. H. Sandstead (1976). Plasma trace metals during total parenteral alimentation. *Gastroenterology 70,* 1022–1025.

Solomons, N. W., R. M. Russell, E. Vinton, A. M. Guerrero, and L. Mejia (1982). Application of a rapid dark adaptation test in children. *J. Pediatr. Gastroenterol. Nutr. 1,* 571–574.

Sommer, A. G., G. Hussaini, Muhilal, I. Tarwotjo, D. Susanto, and J. S. Sarosa (1980). History of night blindness: A simple tool for xerophthalmia screening. *Am. J. Clin. Nutr. 33,* 887–891.

Soustre, Y., M. C. Dop, P. Galan, and S. Hercberg (1986). Dietary determinants of the iron status in menstruating women. *Int. J. Vit. Nutr. Res. 56,* 281–286.

Sowers, M. R., R. B. Wallace, and J. H. Lemke (1985). The association of intakes of vitamin D and calcium with blood pressure among women. *Am. J. Clin. Nutr. 42,* 135–142.

Sowers, M. R., R. B. Wallace, B. W. Hollis, and J. H. Lemke (1986). Parameters related to 25-OH-D levels in a population-based study of women. *Am. J. Clin. Nutr. 43,* 621–628.

Speek, A. J., C. Wongkham, N. Limratana, S. Saowakontha, and W. H. Schreurs (1986). Microdetermination of vitamin A in human plasma using high performance liquid chromatography with fluorescence detection. *J. Chromatog. 382,* 284–289.

Sporn, M. B. and A. B. Roberts (1983). Role of retinoids in differentiation and carcinogenesis. *Cancer Res. 43,* 3034–3040.

Sterner, R. T. and W. R. Price (1973). Restricted riboflavin: Within-subject behavioral effects in humans. *Am. J. Clin. Nutr. 26,* 150–160.

Stryker, S., L. Kaplan, E. A. Stein, M. J. Stampfer, A. Sober, and W. C. Willett (1988). The relation of diet, cigarette smoking, and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. *Am. J. Epidemiol. 127,* 283–296.

Tangney, C. C., R. B. Shekelle, W. Raynor, M. Gale, and E. P. Betz (1987). Intra- and interindividual variation in measurements of β-carotene, retinol, and tocopherols in diet and plasma. *Am. J. Clin. Nutr. 45,* 764–769.

Tarwotjo, I., A. Sommer, K. P. West Jr., E. Djunaedi, L. Mele, B. Hawkins (1987). Influ-

ence of participation on mortality in a randomized trial of vitamin A prophylaxis. *Am. J. Clin. Nutr. 45,* 1466–1471.

Taylor, P. R., M. P. Longnecker, O. A. Levander, S. M. Howe, W. C. Willett, C. Veillon, P. A. McAdam, K. Y. Patterson, J. M. Holden, and M. J. Stampfer (1987). Seasonal variation in selenium (Se) status among free-living persons in South Dakota. *Fed. Proc. 46,* 882.

Thakker, K. M., H. S. Sitren, J. F. Gregory III, G. L. Schmidt, and T. G. Baumgartner (1987). Dosage form and formulation effects on the bioavailability of vitamin E, riboflavin, and vitamin B-6 from multivitamin preparations. *Am. J. Clin. Nutr. 45,* 1472–1479.

Thanassi, J. W., N. T. Meisler, and J. M, Kittler (1985). Vitamin B-6 metabolism and cancer. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* New York, N.Y.: Alan R. Liss, pp. 319–336.

Thomas, L. H., P. R. Jones, J. A. Winter, and M. Smith (1981). Hydrogenated oils and fats: The presence of chemically-modified fatty acids in human adipose tissue. *Am. J. Clin. Nutr. 34,* 877–886.

Thomas, L. H., J. A. Winter, and R. G. Scott (1983). Concentration of 18:1 and 16:1 *trans*-unsaturated fatty acids in the adipose body tissue of decedents dying of ischaemic heart disease compared with controls: Analysis by gas liquid chromatography. *J. Epidemiol. Commun. Health 37,* 16–21.

Thomson, C. D., H. M. Rea, V. M. Doesburg, and M. F. Robinson (1977). Selenium concentrations and glutathione peroxidase activities in whole blood of New Zealand residents. *Br. J. Nutr. 37,* 457–460.

Thomson, C. D., L. K. Ong, and M. F. Robinson (1985). Effects of supplementation with high-selenium wheat bread on selenium, glutathione peroxidase and related enzymes in blood components of New Zealand residents. *Am. J. Clin. Nutr. 41,* 1015–1022.

Thomson, M., M. F. Fulton, D. A. Wood, S. Brown, R. A. Elton, A. Birthwhistle, and M. F. Oliver (1985). A comparison of the nutrient intake of some Scotsmen with dietary recommendations. *Human Nutr. Appl. Nutr. 39,* 443–455.

Thornton, S. P. (1977). A rapid test for dark adaptation. *Ann. Ophthalmol. 9,* 731–734.

Thurnham, D. I. (1972). Influence of glucose-6-phosphate dehydrogenase deficiency on the glutathione reductase test for ariboflavinosis. *Ann. Trop. Med. Parasitol. 66,* 505–508.

Tillotson, J. A. and E. M. Baker (1972). An enzymatic measurement of the riboflavin status in man. *Am. J. Clin. Nutr. 25,* 425–431.

Traber, M. G. and H. J. Kayden (1987). Tocopherol distribution and intracellular localization in human adipose tissue. *Am. J. Clin. Nutr. 46,* 488–495.

Tucker, D. M., H. H. Sandstead, G. M. Logan, L. M. Klevay, J. Mahalko, L. K. Johnson, L. Inman, and G. E. Inglett (1981). Dietary fiber and personality factors as determinants of stool output. *Gastroenterology 81,* 879–883.

Ullrey, D. E. (1981). Selenium in the soil-plant-food chain. In J. E. Spallhotz et al., eds.: *Selenium in Biology and Medicine.* Westport, Conn. AVI Publishing, pp. 176–191.

Underwood, B. A. (1984). Vitamin A in animal and human nutrition. In Sporn, M. B., A. B. Roberts, and D. S. Goodman, eds.: *The Retinoids. Vol. 1.* Orlando, Fla.: Academic Press, pp. 281–392.

Vahlquist, A., G. Michaelsson, and L. Juhlin (1978). Acne treatment with oral zinc and vitamin A: Effects on the serum levels of zinc and retinol binding protein (RBP). *Acta. Derm. Venereol. (Stockh). 58,* 437–442.

van Staveren, W. A., P. Deurenberg, M. B. Katan, J. Burema, L. C. de Groot, and M. D. Hoffmans (1986). Validity of the fatty acid composition of subcutaneous fat tissue

microbiopsies as an estimate of the long-term average fatty acid composition of the diet of separate individuals. *Am. J. Epidemiol. 123,* 455–463.

Versieck, J. (1985). Trace elements in human body fluids and tissues. *C.R.C. Crit. Rev. Clin. Lab. Sci. 22,* 97–184.

Vuilleumier, J. P., H. E. Keller, D. Gysel, and F. Hunziker (1983a). Clinical chemical methods for the routine assessment of the vitamin status in human populations. Part I: The fat-soluble vitamins A and E, and β-carotene. *Int. J. Vit. Nutr. Res. 53,* 265–272.

Vuilleumier, J. P., H. E. Keller, R. Rettenmaier, and F. Hunziker (1983b). Clinical chemical methods for the routine assessment of the vitamin status in human populations. Part II: The water-soluble vitamins B_1, B_2, and B_6. *Int. J. Vit. Nutr. Res. 53,* 359–370.

Wacker, W.E.C. (1980). In *Magnesium and Man,* Cambridge, Mass. Harvard University Press, p. 61.

Wahrendorf, J., A. B. Hanck, N. Munoz, J. P. Vuilleumier, and A. M. Walker (1986). Vitamin measurements in pooled blood samples. *Am. J. Epidemiol. 123,* 544–550.

Wald, N. J., H. S. Cuckle, R. D. Barlow, P. Thompson, K. Nanchahal, R. J. Blow, I. Brown, C. C. Harling, W. J. McCulloch, J. Morgan, et al. (1985). The effect of vitamin A supplementation on serum retinol and retinol binding protein levels. *Cancer Lett. 29,* 203–213.

Wallace, R. B. and R. A. Anderson (1987). Blood lipids, lipid-related measures, and the risk of atherosclerotic cardiovascular disease. *Epidemiol. Rev. 9,* 95–119.

Waring, P. P., D. Fisher, J. McDonnell, E. L. McGown, and H. E. Sauberlich (1982). A continuous-flow (Auto Analyzer II) procedure for measuring erythrocyte transketolase activity. *Clin. Chem. 28,* 2206–2213.

Watson, R. L. and H. G. Langford (1970). Usefulness of overnight urines in population groups. Pilot studies of sodium, potassium and calcium excretion. *Am. J. Clin. Nutr. 23,* 290–304.

Willett, W. C. (1985). Does low vitamin B-6 intake increase the risk of coronary heart disease? In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* New York, N.Y.: Alan R. Liss, pp. 337–346.

Willett, W. C. (1986). Vitamin A and selenium intake in relation to human cancer risk. In Y. Hayashi et al., eds.: *Diet, Nutrition and Cancer.* Japan Sci. Soc. Press, Tokyo/ VNU Sci. Press, Utrecht, pp. 237–245.

Willett, W. (1987). Nutritional epidemiology: Issues and challenges. *Int. J. Epidemiol. 16 (Suppl),* 312–317.

Willett, W. C., M. J. Stampfer, B. A. Underwood, J. O. Taylor, and C. H. Hennekens (1983a). Vitamins A, E, and carotene: effects of supplementation on their plasma levels. *Am. J. Clin. Nutr. 38,* 559–566.

Willett, W. C., M. J. Stampfer, B. A. Underwood, F. E. Speizer, B. Rosner, and C. H. Hennekens (1983b). Validation of a dietary questionnaire with plasma carotenoid and α-tocopherol levels. *Am. J. Clin. Nutr. 38,* 631–639.

Willett, W. C., B. F. Polk, J. S. Morris, M. J. Stampfer, S. Pressel, B. Rosner, J. O. Taylor, K. Schneider, and C. G. Hames (1983c). Prediagnostic serum selenium and risk of cancer. *Lancet ii,* 130–134.

Willett, W. C., M. J. Stampfer, B. A. Underwood, L. A. Sampson, C. H. Hennekens, J. C. Wallingford, L. Cooper, Hsieh C-C, and F. E. Speizer (1984a). Vitamin A supplementation and plasma retinol levels: a randomized trial among women. *J.N.C.I. 73,* 1445–1448.

Willett, W. C., B. F. Polk, B. A. Underwood, M. J. Stampfer, S. Pressel, B. Rosner, J. O. Taylor, K. Schneider, and C. G. Hames (1984b). Relation of serum vitamins A and E and carotenoids to the risk of cancer. *N. Engl. J. Med. 310,* 430–434.

Willett, W. C., B. F. Polk, B. A. Underwood, and C. G. Hames (1984c). Hypertension detection and follow-up program study of serum retinol, retinol-binding protein, total carotenoids, and cancer risk: A summary. *J.N.C.I. 73,* 1459–1462.

Willett, W. C. and B. MacMahon (1984d). Diet and cancer—An overview. *N. Engl. J. Med. 310,* 633–638.

Wittpen, J., S. Tseng, A. Sommer (1986). Detection of early xerophthalmia by impression cytology. *Ach. Opthalmol. 104,* 237–239.

Wood, P. D., P. S. Bachorik, J. J. Albers, et al. (1980). Effects of sample aging on total cholesterol values determined by the automated ferric chloride-sulfuric acid and Liebermann-Burchard procedures. *Clin Chem. 26,* 592–597.

Wood, D. A., R. A. Riemersma, S. Butler, M. Thomson, C. Macintyre, R. A. Elton, and M. F. Oliver (1987). Linoleic and eicosopentaenoic acids in adipose tissue and platelets and risk of coronary heart disease. *Lancet i,* 177–183.

Yamori, Y., Y. Nara, N. Iritani, R. J. Workman, and T. Inagami (1985). Comparison of serum phospholipid fatty acids among fishing and farming Japanese populations and American inlanders. *J. Nutr. Sci. Vitaminol. 31,* 417–422.

Yang, X. Y., H. K. Naito, and R. S. Galen (1986). Urinary calcium, phosphorus, magnesium and copper in 24-hour and random samples from normal subjects. *Trace Elem. Med. 3,* 81–86.

Young, V. R. and P. L. Pellett (1987). Protein intake and requirements with reference to diet and health. *Am. J. Clin. Nutr. 45,* 1323–1343.

10

Anthropometric Measures and Body Composition

Anthropometric variables, particularly weight and height, are probably the most commonly employed measures of nutritional status in epidemiologic studies because of their simplicity and ease of collection. In adults, measures of body dimensions and mass are used to represent directly nutritional status, to compute the absolute size of the major body compartments, such as lean body mass and adipose mass, to estimate relative body composition, such as fatness, and to describe body fat distribution.

This chapter begins with an overview of weight and height, including their relationships to nutritional status, their use in epidemiologic studies, and the reproducibility and validity of these measurements. Next, the concept of major body compartments is discussed and methods of measuring them are considered. The major part of the chapter addresses the assessment of relative body composition, or fatness, using densitometry, combinations of weight and height, skinfold thicknesses, and the newer methods of bioelectric resistance and conductance. Finally, the evaluation of body fat distribution is reviewed.

Throughout, an emphasis is given to methods that are likely to be used by epidemiologists themselves. Methods more suitable for laboratory application are described briefly, as they can provide standards against which field methods can be compared. For each method, attention is given to measurement error. Other reviews of methods for assessing human body composition that may be helpful have been published by Lukaski (1987), Roche and colleagues (1987), and Garrow (1983).

WEIGHT AND HEIGHT

Relationship to Nutritional Status

The voluminous literature regarding the use of weight and height to represent childhood nutritional status is largely related to the assessment of protein–calo-

ric deficiency and is outside the scope of this book. Excellent reviews are available elsewhere (Zerfar, 1979; Garn, 1979). In adults, weight and height are sometimes used directly to represent nutritional status. In populations where food is widely available, such as the United States and most other industrialized countries, it is unlikely that a substantial portion of the population has experienced energy restriction intake sufficient to limit growth. In such countries, height is predominantly determined by genetic factors; however, adult height does reflect nutritional factors in other cultures. Japanese–Americans have attained substantially greater height during successive generations (Insull et al., 1968) and major secular increases in height have been observed within Japan (Hirayama, 1978).

The effect of nutrition on height is complex, as early unconstrained intake will not only accelerate growth, but will also cause early maturity (Frisch and McArthur, 1974) and closure of the epiphyses, resulting in cessation of long bone growth. Because growth of the trunk can continue, sitting height may provide a more sensitive indicator of energy availability before age 20, particularly during adolescence. Biacromial (between the tips of the clavicles) and biiliac diameters also reflect growth during development and thus, potentially, early nutritional factors.

Use of Height and Weight in Epidemiologic Studies

Height and other measures of frame size may be particularly useful in epidemiologic studies because they may reflect an influence of diet in the remote past that may be difficult to measure in any other way. Furthermore, in case-control studies, these measurements are generally not affected by disease or its treatment (except, perhaps, in the case of spinal metastases). For example, Brinkley and co-workers (1971) have used measurements of height and other body dimensions in a case-control study of breast cancer; the positive associations they observed with biacromial and biiliac diameters support the hypothesis that early nutritional status is associated with breast cancer risk. Unfortunately, their data on sitting height (which was inversely related to breast cancer risk) are difficult to interpret as this decreases at older ages due to compression of the vertebrae and they did not account for the older age of the breast cancer cases in their analysis.

Micozzi (1985) has used height as a measure to address the hypothesis proposed by de Waard (1975) that early energy availability might explain some of the international differences in breast cancer risk. In an ecologic study he observed a strong correlation ($r = 0.68$) between average height at age 18 and national age-specific breast cancer incidence.

Limitations of height as a measure of early nutritional status are apparent: some of these measurements may be difficult to obtain in many studies, and in older adults vertebral collapse causes a reduction in height. Because genetic factors strongly influence these dimensions, failure to find associations in epidemiologic studies may simply reflect a lack of variation in energy availability during growth within the study population.

Weight itself is rarely used alone in studies among adults as it is related to both height and body composition. *Change* in weight over time, however, in

effect, controls for overall body size and is thus frequently used as a measure of change in body composition.

Reproducibility and Validity of Weight and Height Measurements

Details of methods to insure optimal precision and accuracy in the collection of weight, height, and other anthropometric variables are discussed elsewhere (Rose et al., 1982; Lohman et al., 1988) and are mentioned only briefly here. The manual by Lohman and colleagues contains detailed instructions with photographs and will probably become the standard for these measurements. Weight is optimally measured on a platform (beam balance) scale with the subject nude or lightly clad and not wearing shoes. Unfortunately, platform scales are often not sufficiently portable for many epidemiologic applications, such as in-home interviews. In this case, it may be necessary to use spring (e.g., bathroom-type) scales. Particular care must be taken to calibrate these scales regularly using standard weights. When direct measurement of standing height is possible, it should be made without shoes, the back square against a metal wall tape, and with eyes looking straight ahead. A set square should rest on the scalp and against the wall tape to read the height, which is recorded to the nearest 0.5 cm (Rose et al., 1982).

In general, body weight is among the most precise biologic measurements even with simple and imperfect conditions. For repeated measurements of weight, within the hour, among children, a standard deviation $\pm$ 0.02 kg has been reported by Habicht and colleagues (1979). If a single measurement is used to represent weight over a longer period, such as months, factors such as changes in hydration, recency of food intake, and intended weight change will increase error substantially more than that due to the technical aspect of weighing. The net effect of both biologic and technical sources of variation can be evaluated by repeated measurements over a period of months. For example, we found a correlation of 0.96 for repeated assessments of weight at an interval of approximately 1 year among adult women, even though the first value was based on self-report and the other on measurement using a bathroom-type scale. This degree of reproducibility is far greater than for most biochemical or physiologic measurements and indicates that, with reasonable care, imprecision in measurement of weight is not likely to be a serious issue in most epidemiologic studies. In studies that involve *change* in weight over months or a few years, precision in measurement is much more critical as measurement errors contribute twice (at the beginning and the end) and the magnitude of weight changes is usually small compared with differences in attained weight between persons.

Self-reported Weight and Height

In many epidemiologic studies weight and height are necessarily based on self-reports. The validity of these self-reports has been addressed by several investigators (Stunkard and Albain, 1981; Palta et al., 1982; Stewart 1982; Stewart et al., 1987). Stunkard and Albain found that reported weight averaged slightly lower (2 to 6 lbs) than directly measured weight, but the correlation between

them was 0.99 among U.S. men and women. These authors, however, found that self-report among Danish men and women was less valid (r = 0.91 to 0.97 for groups defined by age and sex). Stewart (1982) observed correlations of 0.98 or more among both U.S. men and women when reported and measured weights were compared. Palta and co-workers (1982) also found that errors were small among a large group of U.S. adults for self-reported weight and height when the report was compared with direct measurements. Weight was understated by 1.6 percent among men and by 3.1 percent among women, whereas height was overstated by 1.3 percent by men and 0.6 percent by women. Stewart and colleagues (1987) observed similar small systematic tendencies to underreport weight and overreport height among 1598 subjects in New Zealand; the correlation between reported and direct measures was 0.96 for height, 0.98 for weight, and 0.94 for relative weight. These authors point out that, due to the small tendencies to underreport weight and overreport height, relative weight based on these measures will be biased downward. This causes some degree of systematic misclassification when comparing data on an absolute scale, such as weight/height^2 (often called body mass index or Quetelet's index). Epidemiologic measures of association, such as relative risks, however, will not be appreciably affected by this degree of measurement error (see Chapter 12).

In a study among 1805 Japanese men living in Honolulu, Rhoads and Kagan (1983) compared the report of weight at age 25 recalled 20 to 30 years later with actual weights at age 25 recorded at the time of draft registration. Recalled weight averaged 2.2 percent higher than that recorded at registration, and the correlation between estimates was 0.80.

In summary, even self-reported weight and height appear to be sufficiently accurate and precise in most circumstances so that their errors have minimal effect on epidemiologic measures of association. A more serious degree of error arises when combinations of weight and height are used to represent body fat composition (discussed later).

Measurement of Major Body Compartments

The human body has frequently been considered as two major compartments: adipose (storage fat) and lean tissue. Conceptually, lean mass is involved in highly active metabolic processes, so that nutritional requirements are primarily related to the size of this compartment. Adipose, in contrast, has traditionally been regarded as metabolically inactive and its main energy requirement is for transportation, as cargo, from place to place. This concept, although serving as a useful first approximation in many instances, is quite incomplete. A somewhat more detailed scheme is described later.

Body compartments are most often of interest as determinants of disease when considered in relation to overall body size. For example, it is likely that a large absolute fat mass will have more adverse metabolic effects in a person whose other body compartments are small than in a person who is large in all respects. Similarly, a small absolute bone mass will be more likely to fracture in a large person than in a small person. For some methods of determining relative body composition, such as percent body fat, it is first necessary to measure abso-

lute compartments. The absolute size of compartments may also be inherently interesting. In particular, effects of absolute intakes of specific nutritional factors are probably important in relation to the absolute size of body compartments rather than to their relative size.

Adipose is tissue primarily composed of storage fat, mainly in the form of triglyceride. Although certainly less metabolically active in terms of energy and nutrient requirements, adipose tissue does play an important role in hormone metabolism, such as in the synthesis of estrogen in postmenopausal women (Grodin et al., 1973). The major fat stores are under the skin and in the abdomen, although considerable amounts can also be located within muscles and around other organs such as the heart and kidneys.

Lean body mass is considered to be the body component that is not adipose, and is, therefore, more metabolically active. Lean body mass is, of course, extremely heterogeneous as it includes bone, muscle, extracellular water, nervous tissue, and all cells other than adipocytes (Fig. 10–1). If the lean body mass (weight) is known, the adipose component can be calculated as the difference between the lean and total body mass.

The division of the body into adipose and lean body mass provides a useful anatomic distinction. Adipose tissue, however, does not include all fat in the body, such as the lipid contained within other cells, hepatocytes, or important structural lipids in cell membranes or the nervous system. Thus, it is sometimes convenient to partition the body on the basis of chemical composition, rather than anatomy, between fat mass and fat-free mass components. These compartments correspond closely, but not exactly, to adipose and lean body mass for the reasons noted previously. The distinction, however, does not appear to be of major practical importance as lipid comprises only about 2 percent of the lean

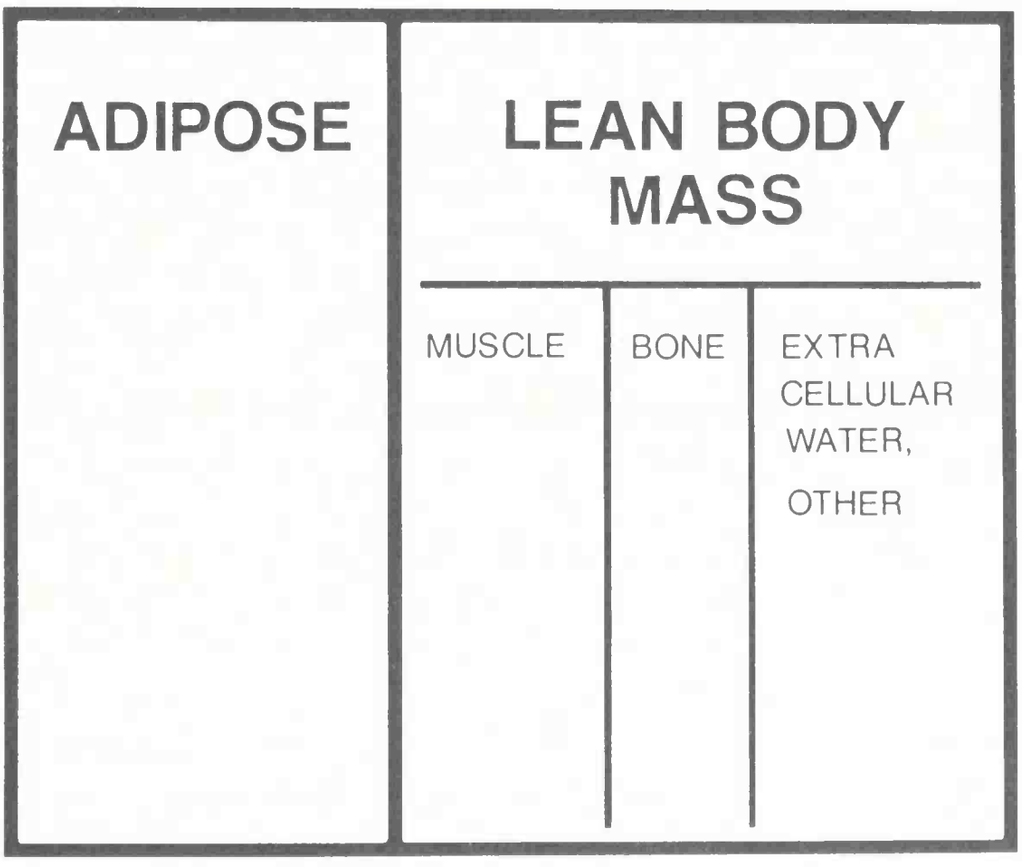


Figure 10–1. Components of total body mass.

body mass (Sheng and Huggins, 1979). At present, few practical methods exist to measure the adipose or total fat mass components directly; indirect methods are discussed later.

The measurement of lean body mass usually involves two assumptions: that water is distributed only to this component (as the water content of adipose tissue is very low) and that the concentration of water in the lean body mass is constant among persons. Based primarily on data from animals, it has been estimated that the fraction of water in the lean body mass is 0.732. Thus, lean body mass = total body water/0.732. The limitations of using total body water to estimate lean body mass (and indirectly, body fat composition) have been reviewed by Sheng and Huggins (1979). They note that the assumption of a constant proportion of water in the fat-free mass cannot be perfectly correct, as this varies depending on the state of hydration and the relative subcomponents of this mass. For example, bone mass that contains relatively little water but is quite dense, can vary considerably among postmenopausal women.

Measurement of total body water is usually accomplished by dilution; a known amount of a substance that is distributed to the total body water is administered and, after sufficient time for "mixing," the concentration of this substance in the body water is measured. The total amount administered is then divided by the concentration, providing a measure of total body water. Deuterium oxide (heavy water, derived from a stable isotope of hydrogen) is commonly employed for this determination (Lukaski and Johnson, 1985); tritium oxide and antipyrine have also been used. Dilution methods are sufficiently simple and inexpensive to be practical in many epidemiologic settings: an oral dose is administered and urine or saliva samples are collected over the next 2 to 4 hours. To minimize error, subjects should be fully hydrated but not edematous; this is usually not problematic in healthy adult populations.

Total body water (and thus lean body mass) has been extensively studied in relation to anthropometric measurements. Although the human body is three-dimensional, lean body mass appears to have a direct linear relationship with height among adults rather than an exponential relationship; presumably the adult body is closer to a straight line than a sphere (Fig. 10–2). Thus, height can be used as a first approximation to represent lean body mass in epidemiologic analyses. Lean body mass, however, is correlated with weight even among people of the same height, both because lean mass itself has weight and because adiposity is positively correlated with lean body mass (Roche, 1984). Thus, improved estimates of lean body mass can be calculated using measures of both height and weight. Such equations have been provided by Watson and colleagues (1980) who have also summarized equations derived from other populations. These authors found that calculations of total body water based on age, height, and weight were highly correlated with measurements by dilution ($r = 0.84$ for men and 0.86 for women); using exponential terms of weight and height did not improve these correlations. The equations predicting total body water (liters) from age (years), height (cm) and weight (kg) were:

$$\text{Men: Total body water} = 2.447 - 0.09516\ \text{age} + 0.1074\ \text{height} + 0.3362\ \text{weight}$$

$$\text{Women: Total body water} = -2.097 + 0.1069\ \text{height} + 0.2466\ \text{weight}.$$

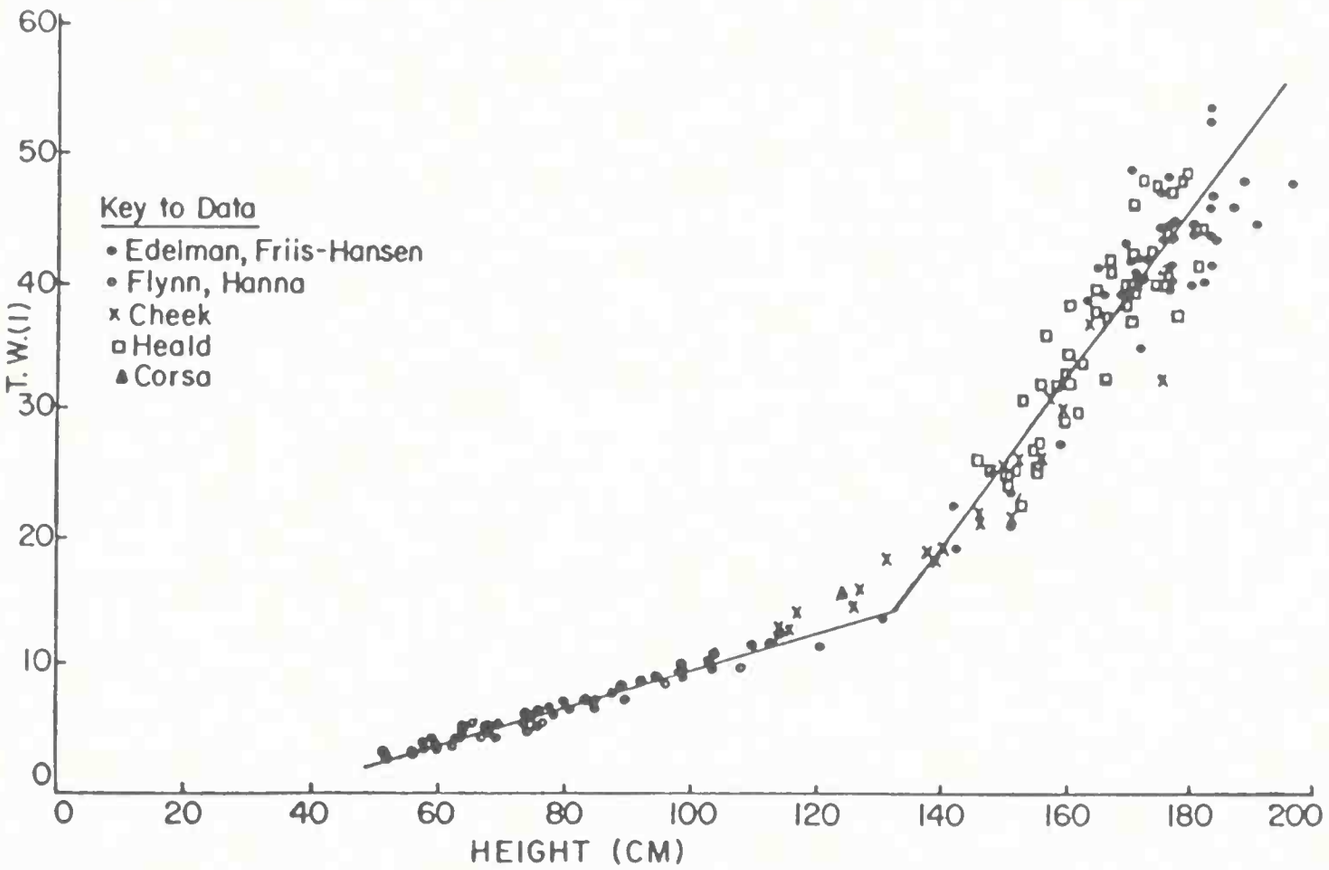


Figure 10–2. Total body water in relation to height (from Mellits and Cheek, 1970, reproduced with permission).

As noted, age appeared to be an important variable for men, but not for women.

Body cell mass represents the part of lean body mass composed of cells, which is mainly muscle, and excludes extracellular water and bone minerals (Moore et al., 1963). Although less frequently used because it is difficult to measure, this is the component of lean body mass that is most active with respect to energy utilization. Measurement of active cell mass is usually based on the determination of total body potassium, using the fact that body potassium is largely intracellular. This measurement is best accomplished by counting the gamma-ray emissions of potassium-40, a naturally occurring radioisotope. Because this method measures a natural isotope (which occurs as a known proportion of all potassium isotopes), no administration of a tracer is required; however, the method requires a special, generally unavailable, counting chamber (Lukaski, 1987).

Bone mass is the other major component of lean body mass. Bone mass measurement can enhance the accuracy of other body components that are measured indirectly, and is of inherent interest in the study of fractures. Photon absorptiometry is most commonly used to measure bone mass and is based on the principle that the mineral content of the bone being studied is directly proportional to the absorbed energy from a photon beam emitted by a radionucleide. Lukaski (1987) provides a detailed review of this technology and its accuracy. Single photon absorptiometry uses a simple and relatively inexpensive device to measure bone mass or density in the arm or legs; a site over the distal radius is the most commonly used. Although highly reproducible, even multiple measurements do not provide an accurate assessment of total bone mass, in part

because much of the bone mass is contained in the axial skeleton and the densities at various sites are poorly correlated.

Dual photon absorptiometry uses the differential absorption of photon beams with two distinct energy levels to differentiate between bone mineral and soft tissue mass. Coupled with a body scanning device, this provides an accurate measurement of total body bone mass as well as lean and fat masses. Although the radiation exposure is sufficiently low to make this appropriate for some research applications, the instrument is relatively expensive. Practical limitations are likely to limit its application in epidemiologic studies.

MEASUREMENT OF RELATIVE BODY COMPOSITION

Introduction

Because of widespread interest in the health effects of obesity, the most frequent use of anthropometry in epidemiologic studies is to estimate adiposity or percent body fat. In epidemiologic studies, the most commonly used methods to estimate relative body composition are combinations of weight and height, and skinfold thickness. Recently, methods based on electrical resistance and impedance have become available. Because these methods measure adiposity indirectly, it is crucial to consider the degree of error associated with their use, as well as their feasibility. Because densitometry is the generally accepted standard for measuring the percentage of body weight that is fat, this method is described first, even though it is impractical for most epidemiologic applications.

Densitometry

Densitometry (also called hydrostatic weighing) is based on the principle that fat tissue is lighter than fat-free tissue. The ratio of weights measured in air and under water, therefore, provides an estimate of the proportion of total body mass that is composed of fat. In the most widely used technique, subjects wearing a swimming suit are submerged seated on a scale in a tank of water (with a known weight strapped to their body so that they do not float). Because the air in the lungs influences weight under water, residual lung volume is measured by having subjects breath through a snorkel into a special device. Formulas for the calculation of percentage body fat from these data have been developed by Siri (1961) and Brozek and colleagues (1963). Both biologic variation in the density of fat and lean body mass, as well as technical variation in the measurement of density, contribute to error in estimating body fat composition by densitometry. Of these factors, variation in the water content of the lean body mass, in bone size, and the density of bone appear to be the primary source of error and may lead to errors of 3 to 4 percent in predicting body fatness (Lohman, 1981; Lukaski, 1987). Within a population of similar age, sex, and race, however, the biologic sources of error should be considerably less important than in the heterogeneous groups in which sources of error have been evaluated. For demographically homogeneous groups, the magnitude of error associated with densitometry is not well defined.

Combinations of Weight and Height

Weight and height are the most commonly available anthropometric measurements in epidemiologic settings; the literature regarding methods of combining them to best represent adiposity is enormous. The criteria usually employed are (1) that the index should be highly correlated with percent body fat and (2) that the index be uncorrelated with height. The first criterion is obviously most important, but also the most difficult to evaluate as a perfect standard for adiposity is not available and the best methods, such as densitometry, are difficult for practical reasons. More attention has been focused on the second criterion, probably because this is far easier to evaluate. This criterion, however, has become less important with the advent of computers because many multivariate procedures are widely available that can easily provide a statistical adjustment for height (discussed later). If height is not associated with the disease being investigated, the second criterion is largely irrelevant in that setting. Moreover, the second criterion makes the implicit assumption that adiposity is unrelated to height. This appears to be generally true for adults, but is not for children; before puberty, obese children tend to be taller than lean children (discussed later and in Roche, 1984). The two most commonly employed measures of obesity are relative weight (a standardized ratio) and indices of weight and height that are not related to a standard.

Relative weight is the ratio of a subject's observed weight to a standard or expected weight; this may also be expressed as a percentage above or below the standard. The standard weights are frequently derived from a large group of persons of the same height, sex, and (sometimes) age. These may be obtained from an external population, such as the widely used Metropolitan Life Insurance "desirable weights" (Metropolitan Life Insurance Co., 1959); these are based on associations with minimal mortality among insurees, and are revised periodically. In some large investigations, such as the American Cancer Society cohort of 750,000 men and women (Lew and Garfinkel, 1979), the average weights for study participants of the same height, sex, and age are used as standards.

The use of relative weight provides a readily interpretable measure; to say that a group of subjects was 140 to 150 percent of the average weight for their age, sex, and height conveys a meaningful image to almost any reader. The distinct disadvantage of this approach is that findings from different studies are difficult to compare as a wide variety of standards may have been employed. It is not often appreciated, for example, that the Metropolitan Life standards are substantially below the standards based on average weights in other U.S. studies (Manson et al., 1987). Differences in standards are likely to be even greater internationally.

Obesity indices are combinations of weight and height that are not related to a standard. More than 100 years ago, Quetelet (1869) pointed out that weight/height^2 (sometimes called body mass index) was minimally correlated with height. Other investigators have advocated the use of weight/height, weight/height$^{1.5}$, and weight/height^3. Collectively, these have been called power indices. Benn (1971) has advocated the use of an empirically fit value for the exponent of height (p) based on the specific population being studied so that, by definition,

Table 10–1. Correlation coefficients between height and selected indices of obesity (Pearson r)

Source	Subjects	Wt	Wt/Ht	Wt/Ht^2	Wt/Ht^3	$Wt/\overline{Wt}$	Log skin-fold	Body fat (densitometry)
Allen[a] et al. (1956)	55 men	—	0.41	0.16	−0.16	—	—	0.03
	26 women		0.27	0.03	−0.24	—	—	0.05
Khosla et al. (1967)	5000 men (15–64 yrs)	0.43/0.59	0.19/0.37	−0.10/0.08	−0.36/−0.24	—	—	—
Evans and Prior (1969)	432 men	0.44	0.25	0.02	−0.21	—	−0.13[b]/0.15[b]	—
	378 women	0.55	0.33	0.05	−0.18		−0.06[b]/0.19[b]	—
Florey (1970)	1723 men	0.48	—	—	—	—	−0.12(trcps) 0.04(infscp) 0.02(trcps) −0.08(infscp)	—
	2202 women	0.25	—	—	—	—	−0.08(infscp)	
Keys et al. (1972)	180 students	0.46	0.25	0.02	−0.24[c]	0.04	0.04	—
	249 excutives	0.40	0.18	0.06	−0.30[c]	0.01	0.02	—
Goldbourt and Medalie (1974)	9475 men	0.52	0.28	−0.03	−0.31[c]	—	0.09(trcps) 0.05(infscp)	—

Womersley and Durnin (1977)	245 men	—	0.01	−0.22	−0.43[c]	−0.08	−0.13	−0.22
	324 women	—	0.06	−0.10	−0.26[c]	−0.04	−0.06	−0.13
Revicki and Israel (1986)	474 men	0.19	−0.03	−0.24	—	0.01	0.02	—
Killeen et al. (1978)	13,867 children 6–17 yrs (by race and sex)	0.21/0.71	−0.01/0.38[d]	−0.36/0.18	—	—	0.02 0.34(infscp)	—
Michielutte et al. (1984)	832 boys	—	0.76	0.44	−0.27	—	—	—
	836 girls age 5–12 yr	—	0.78	0.50	−0.18	—	—	—
Micozzi et al. (1986)	5808 men (25–74 yr)	0.42 0.21	0.08/0.24 ~0.00/0.08	~0.00/ −0.08	−0.22/−0.33 −0.25/−0.31	—	—	—
	8592 women (25–74 yr)			~0.00/ −0.13		—	—	—

[a]Calculations from Wormersley et al., 1977.

[b]Sum of triceps and infrascapular.

[c]Correlations are for cube root of weight divided by height (ponderal index).

[d]Strongest positive correlations for youngest children.

ht = height, infscp = infrascapular, trcps = triceps, wt = weight.

the index (weight/height^n) is uncorrelated with height. He has further shown that such an index is perfectly correlated with relative weight based on a standard from within the same population. The use of these obesity indices, with the exception of the Benn index, has the considerable advantage that they provide measurement scales that do not vary from study to study, thus facilitating the comparison of findings. The relative merits of the different indices should be considered on the basis of their relationships with true adiposity and, to a lesser degree, on being uncorrelated with height.

Multivariate adjustment for height provides a simple alternative to the use of relative weight or obesity indices. Weight and height can both be entered as independent variables in a multiple regression model predicting the outcome of interest; this provides a measure of the effect of weight independent of height, thus by definition weight uncorrelated with height. Conceptually, this can be thought of as the effect of weight among individuals of identical height.

Although the meaning of weight in this multivariate model containing weight and height is relatively clear, the converse is not; the interpretation of height adjusted for weight is conceptually unclear and of little interest as it is strongly related to body composition. (This was appreciated by the wit of yore who commented that he was not fat, just short for his weight.) If both height and weight-adjusted-for-height (an estimate of obesity) are of interest, a two-step procedure analogous to that suggested for adjusting nutrient intake for total caloric intake can be employed (see Chapter 11). First, a simple regression model is used with height as the independent variable (x) and weight as the dependent variable (y). The resulting residual of weight on height provides, by definition, a measure of weight uncorrelated with height. (This measure thus has all the advantages of the Benn index, with the added feature of having the usual scale of weight, i.e., kilograms or pounds.) In the second step, height and the residual of weight on height can both be entered in the multivariate model. This simultaneously provides the full effect of height as well as a measure of weight independent of height.

The assumption that true adiposity is unrelated to height is central to the interpretation of relative weight and obesity indices. This relationship has been addressed by examining the correlations between height and adiposity measured by skin folds or by densitometry (Table 10–1). In adults, it does appear that the correlation between height and adiposity is minimal; in only the data of Womersley and Durnin (1977) does there appear to be a slight inverse relationship.

Among the obesity indices shown in Table 10–1, the correlation with height has generally been lowest for Quetelet's index (weight/height^2). The Benn index, using an empirically fit exponent for weight (Benn, 1971; Lee et al., 1981) does not seem to provide any clear advantage as the correlation between height and weight/height^2 is typically already small (Garn and Pesick, 1982; Colliver et al., 1983). Others have found that the exponent of 1.5 produces a slighty lower correlation with height among women (Micozzi et al., 1986).

Relationships of height and weight among children are considerably more complex, probably because this is a period of active growth; an exhaustive review is beyond the scope of this book. At some ages, however, height is positively associated with adiposity to a degree such that an assumption of indepen-

Table 10–2. Correlation of skinfold measures with anthropometric indices of obesity (Pearson r)

Source	Subjects	Skinfold	Ht	Wt	Wt/Ht	Wt/Ht^2	Wt/Ht^3
Flory (1970)	1723 men	triceps	0.12	0.44	0.45	0.42	0.36[a]
	1723 men	infrascapular	0.04	0.59	0.64	0.64	0.59[a]
	2202 women	triceps	0.02	0.47	0.47	0.46	0.44[a]
	2202 women	infrascapular	0.08	0.61	0.65	0.66	0.64[a]
Keys et al. (1972)	180 students 18–24 yr	Sum of triceps + infrascapular	0.06	0.78	0.83	0.85	0.81[a]
	249 executives		0.00	0.72	0.77	0.78	0.74[a]
Goldbourt and Medalie (1974)	9,475 Israeli men	triceps	0.09	0.39	0.42	0.40	0.36[a]
		infrascapular	0.05	0.54	0.60	0.60	0.56[a]
Killen et al. (1978)	13,687 children 6–17 yr by age and sex	infrascapular	—	—	0.55–0.81	0.61–0.83 [b]	0.47–0.81
Michielutte et al. (1984)	832 boys 5–12 yr	triceps	—	—	0.73	0.81[b]	0.69
	835 girls 5–12 yr	triceps	—	—	0.73	0.81[b]	0.64
Revicki and Israel (1986)	474 men	7 skinfolds (computed % fat)	0.01	0.71 0.70[c]	0.75 0.74[c]	0.76 0.72[c]	0.73 0.72[c]
Micozzi et al. (1986)	5808 men	infrascapular	~0.00	0.69	0.75	0.77	0.74
	8592 women	infrascapular	−0.09	0.76	0.79	0.80	0.79

[a]Correlations for cube root of weight divided by height (ponderal index).

[b]Correlations lowest for younger children.

[c]Age-adjusted.

ht = height, wt = weight.

dence is materially violated (Killeen et al., 1978). In this situation one could use the combination of weight and height most strongly correlated with obesity and adjust for height, if it is related to the outcome being studied, with multivariate analysis. Alternatively, the use of a more direct measure, such as skinfolds, may be preferable.

The validity of combinations of weight and height as measures of adiposity has frequently been assessed by correlating these with skinfold thickness (Table 10–2). Because skinfold thicknesses themselves are imperfect indicators of adiposity, the absolute value of these correlations should not be interpreted as a direct measure of validity. Comparing the correlations of skinfolds with different combinations of weight and height, however, may be useful to assess their relative degrees of validity. As seen in Table 10–2, the correlations with skinfolds are quite similar whether one uses weight/height or $weight/height^2$, and the correlations with $weight/height^3$ are only slightly reduced. Indeed, the use of weight alone is nearly as good as any weight index. This result is not surprising as it can be readily appreciated that most variation in weight between individuals is independent of height as adult heights do not vary dramatically; within one age and

sex group the range from the tallest to the shortest is typically only about 20 percent.

The Influence of Frame Size

It is commonly assumed that weight should be evaluated in relation to frame size and that an accurate measurement of skeletal dimensions may improve the validity of obesity indices that are based simply on weight and height. Indeed, height itself is basically a one-dimensional measure of frame size that is easily available in most studies. Widely used "ideal weights," such as those published by Metropolitan Life, are often provided for small, medium, and large frame sizes. These categories of frame size, however, have no quantitative definition and are left to individual judgment.

In addition to height, other measures of skeletal dimensions include biacromial diameter, knee and elbow width (Frisancho, 1984), biiliac diameter, and chest depth (Garn et al., 1986); such measurements have sometimes been combined into indices of frame size (Katch and Freedson, 1982). Katch and colleagues (1982) have demonstrated that frame sizes based on self-report or a subjective rating by an expert correspond poorly with a standardized measurement of frame size. Roche (1984) has reviewed studies that address the issue of whether measures of frame size improve the prediction of body fat composition above and beyond that provided by simple weight and height. Overall, there appears to be no consistent evidence that frame size measurements in addition to height provide any important refinement in the estimation of obesity. This is probably expected because, as demonstrated by the data in Tables 10–2 and 10–3, even height provides only a modest incremental improvement in prediction of body fat composition; further refinements in frame size estimation are likely to produce smaller marginal gains. Although additional work is warranted to identify simple measures of frame size that may improve the interpretation of weight, the cost and difficulty involved in obtaining such measurements are unlikely to be justified in epidemiologic studies of obesity.

Skinfold Measurements

Next to combinations of height and weight, skinfolds are probably the most widely used method to measure body composition in epidemiologic studies. This method has conceptual appeal because it provides a direct measure of body fat; its major limitations are that not all fat is accessible to the calipers (such as intraabdominal and intramuscular fat), and that the distribution of subcutaneous fat can vary considerably over the body. This variability in distribution of subcutaneous fat creates difficulties when measurements at one or only a few sites are used to represent overall body fat composition; however, these distributions may be of interest in their own right (discussed later).

The technical aspects of skinfold measurements and considerations for selecting specific sites are discussed in detail elsewhere (Habicht et al., 1979; Lohman, 1981; Mueller and Stallones, 1981; Rose et al., 1982, Roche, 1984). In general, skinfold measurements are substantially less reproducible than most

Table 10–3. Correlation coefficients between densitometry estimates of body fat composition and anthropometric indices of obesity (Pearson r)

Source	Subjects	Wt	Wt/Ht	Wt/Ht^2	Wt/Ht^3	Skinfold
Allen et al. (1956)	55 men	—	0.70	0.72	0.68	—
	26 women	—	0.74	0.80	0.77	—
Parizkova (1961)	62 normal adolescents	—	—	—	—	0.74 (triceps) 0.80 (infrascapular)
Seltzer et al. (1965)	32 obese adolescent girls	—	—	—	—	0.69 (triceps) 0.59 (infrascapular)
Keys et al. (1972)	180 students	—	0.83	0.85	0.79[b]	0.85[c]
	249 executives	—	0.66	0.67	0.66[b]	0.82[c]
Womersley and Durnin (1977)	245 men	—	0.68	0.71	0.72[b]	0.84
	324 women	—	0.81	0.82	0.84[b]	0.86
Harsha, Frerichs, and Berenson (1978)	242 black and white children	0.32	—	—	—	0.81 (triceps) 0.76 (infrascapular)
Roche et al. (1981)	68 boys (6–12 yr)	0.33	0.73	0.68	0.74	0.84 (triceps) 0.74 (infrascapular)
	49 girls (6–12 yr)	0.23	0.69	0.55	0.62	0.83 (triceps) 0.68 (infrascapular)
	63 boys (13–17 yr)	0.30	0.71	0.61	0.71	0.78 (infrascapular) 0.72 (infrascapular)
	81 girls (13–17 yr)	0.72	0.77	0.77	0.74	0.83 (triceps) 0.81 (infrascapular)
	141 men (18–49 yr)	0.67	0.64	0.77	0.75	0.70 (triceps) 0.75 (infrascapular)
	135 women (18–49 yr)	0.70	0.69	0.76	0.75	0.77 (triceps) 0.71 (infrascapular)
Revicki and Israel (1986)	474 men	0.66	0.70	0.71	0.69	0.84 (7 measures)
		—	0.52[d]	0.58[d]	0.58[d]	

[a]Calculations from Wormersley et al., 1977.

[b]Correlations are for cube root of weight divided by height (ponderal index).

[c]Sum of triceps and infrascapular.

[d]Age-adjusted.

other anthropometric measures, such as weight, height, and limb and girth circumferences (Bray et al., 1978; Habicht et al., 1979; Lukaski, 1987).

Ruiz and colleagues (1971) formally investigated sources of variation in skinfold thickness measurements. They found that a small difference (2.5 cm) in the site of measuring the triceps skinfold, for example, resulted in a difference as large as 50 percent in the average skinfold. Other factors, such as the manner in which the skinfold was picked up and the depth of caliper bite, contributed less to variation. Jointly, these factors contribute to the substantial interobserver variation that has typically been reported for such measurements. Because of this relatively high degree of error variance, skinfold thickness measurements are of limited use in following changes in obesity over time (Bray et al., 1978).

The validity of skinfold thickness assessed by calipers as a measure of true

subcutaneous adipose thickness (as opposed to being a measure of body fat composition, which will be discussed next) has been assessed by comparing data obtained by calipers and by computed tomography or ultrasound. Fanelli and Kuczmarski (1984) and Kuczmarski and colleagues (1987) found that subcutaneous fat measurements by ultrasound were not superior to skinfold measurements by caliper in predicting body fat composition determined by densitometry among relatively lean individuals. Among obese adults, however, the ultrasonic measurements proved to be superior. Among the obese group, correlations between skinfold thickness and ultrasonic measurements at the same site ranged from 0.30 (waist) to 0.72 (thigh and biceps). Seidell and colleagues (1987) compared the sum of paraumbilical and suprailiac skinfolds with the cross-sectional area of subcutaneous fat measure by computed tomography. They observed high correlations for both men (0.83) and women (0.88).

It is possible that ultrasound measurements may become commonly used in epidemiologic studies, as the method is simple and potentially safe. It should be noted that many of the same limitations of the traditional skinfold technique (sensitivity to the exact placement of the device and the general variation of subcutaneous fat over the body) also applies to the ultrasound method.

Validity of Relative Weights, Obesity Indices, and Skinfold Thicknesses as Measures of Body Fat Composition

The validity of epidemiologic measures of body composition can be assessed by comparison with more accurate and precise methods. Because even the best methods are indirect, the choice of an optimal "gold standard," as for dietary intake, is not completely clear. In addition to being highly accurate, it would be desirable that any error associated with the gold standard be independent of error in the method being evaluated so that correlation does not occur simply on the basis of errors that are common to both approaches. Until the present, most studies of validity have employed densitometry (underwater weighing) as the standard method. Although this method is not perfect due, for example, to variation in the bone density of subjects, it is a reasonable choice as any errors should be independent, and it has been in widespread use for decades. It would be reassuring if several of the more sophisticated approaches for measuring obesity (e.g., densitometry, deuterium dilution, and electrical conductance) were compared with each other; if very high correlations were observed between them, say on the order of 0.95, this would provide reassurance that they were all providing similarly precise information and could equally serve as standards. A number of studies in which obesity indices and skinfold thicknesses have been compared with densitometry measurements are summarized in Table 10–3. Correlations with obesity indices have ranged from approximately 0.60 to 0.85. Although the correlations with weight/height, weight/height^2 and weight/height^3 are similar, Quetelet's index (weight/height^2) tends to be slightly more strongly correlated. The Benn index (weight/height^p) had no higher correlation with densitometry than weight/height^2 in a large study among men (Revicki and Israel, 1986). The correlations of skinfold thicknesses with densitometry have a similar range of coefficients and are not clearly higher than for the obesity indices; it is

possible that the use of multiple skin folds may improve the correlation with densitometry.

Using another approach to evaluate the relative validity of different epidemiologic measures of obesity, Criqui and co-workers (1982) compared various indices with blood triglyceride, total cholesterol, blood pressure, and fasting glucose (which are all thought to be related to obesity) among a large population of men and women (Table 10–4). For each outcome, weight/height^2 and relative weight exhibited the strongest correlations, providing evidence that these are the most biologically relevant measures of obesity.

Although the correlations with densitometry seen in Table 10–3 are reasonably high, there are several reasons to believe that they overrepresent the validity of obesity indices and skinfolds in the context of most epidemiologic studies. Because the magnitude of a correlation coefficient is directly related to the degree of variation in the parameter being studied, in this case the between-person variation in obesity, the observed correlation coefficient is applicable only to study populations with a similar variation in obesity. In the published studies, it is often unclear how subjects were selected; however, it seems that they were frequently enriched with an atypically high representation of obese subjects. This would tend to lead to higher correlations than would be observed, given the same degree of accuracy, in a general population. Furthermore, correlations have not been adjusted for age in most published reports. This also overstates the relevant variation in obesity because obesity tends to be strongly correlated with age and virtually all epidemiologic analyses adjust for age. In a study that examined the association of weight-for-height indices with obesity measured by densitometry, Womersley and Durnin (1977) found a correlation of 0.71 for weight/height^2 among men when all ages were combined, but correlations ranging from 0.49 to 0.62 within specific 10-year age groups. Correlations with weight/height^2 were somewhat higher among women, being 0.81 overall, and ranging from 0.64 to 0.91 within specific age groups. Similarly, in a study among 474 men (with unclear basis for selection), Revicki and Israel (1986) found that the correlation between weight/height^2 decreased from 0.71 to 0.58 with adjustment for age. For these reasons, it is difficult to determine the true degree of validity for measures such as Quetelet's index in the context of epidemiologic studies on the basis of published data. It seems likely, however, that the correlation with true percent body fat composition in general populations is likely to be on the order of 0.5 or 0.6 for men and perhaps slightly higher for women. On the basis of a vast literature in which obesity indices have been predictive of a wide variety of health outcomes, it is clear that such indices are extremely useful. As shown in Chapter 12, however, this degree of error in measuring obesity suggests that the associations of obesity with disease are likely to have been substantially underestimated, perhaps by a factor of two.

In summary, on the basis of the previous data, one or two skinfold measurements or any of the obesity indices provide approximately similar estimates of relative body fat composition. Among the obesity indices, weight/height^2 appears at least as good as the others as a measure of relative adiposity and is usually optimal with respect to lack of correlation with height. Although the use of other exponents for height may slightly reduce the correlation with height in

Table 10–4. Correlations of height (ht), weight (wt), and obesity indices with risk factors in men and women aged 20–79 years, Rancho Bernardo, California, 1972–1974[a]

	Men aged 20–79 yr (n = 2266)							Women aged 20–79 yr (n = 2690)						
	Weight	Height	Wt/Ht	Wt/Ht^2	$\sqrt[3]{Wt/Ht}$	$-Ht/\sqrt[3]{Wt}$	Relative weight	Wt	Ht	Wt/Ht	Wt/Ht^2	$\sqrt[3]{Wt/Ht}$	$-Ht/\sqrt[3]{Wt}$	Relative weight
Age	−0.16	−0.29	−0.09	0.00	−0.14	−0.14	−0.01	0.04	−0.29	0.11	0.19	0.06	0.07	0.17
Cholesterol	0.01	−0.08	0.04	0.07	0.02	0.03	0.07	0.06	−0.17	0.10	0.15	0.07	0.09	0.14
Log triglyceride	0.21	0.01	0.23	0.24	0.22	0.22	0.25	0.19	−0.10	0.23	0.25	0.21	0.20	0.25
Systolic blood pressure	0.01	−0.17	0.07	0.12	0.03	0.02	0.12	0.14	−0.20	0.20	0.25	0.16	0.15	0.25
Diastolic blood pressure	0.14	−0.04	0.17	0.19	0.15	0.15	0.19	0.19	−0.09	0.22	0.24	0.20	0.19	0.25
Fasting plasma glucose	0.08	−0.03	0.10	0.12	0.09	0.09	0.12	0.10	−0.06	0.12	0.13	0.10	0.10	0.13

[a]All correlations of absolute magnitude, 0.07 in women and 0.08 in men, or greater, are significant at $p < 0.001$.

From Criqui et al., 1982.

a particular population, this rarely outweighs the substantial advantages in comparability among studies that are associated with the use of weight/height^2. Keys and colleagues (1972) have concluded similarly that weight/height^2 is the preferable measure of relative weight in epidemiologic studies. Nevertheless, if height is strongly associated with disease in a particular study, it is important to be certain that the obesity index used is not associated with height. If so, any of a variety of multivariate methods can be used to control for confounding due to height.

The validity of Quetelet's index and other obesity indices as measures of relative body composition, represented by the correlations in Table 10–3, is certainly less than perfect. As discussed earlier, this is not the result of error in measuring weight or height; the primary source of error is that these indices reflect the weight of both lean body mass and fat tissue. Bone mass, as indicated by frame size (Garn et al., 1986), and active cell mass both contribute to the correlation of lean body mass with obesity indices based on weight and height. It is, therefore, important to consider that associations between obesity indices and other variables can be due to differences in lean body mass as well as adiposity. This potential for confounding, however, is due not only to the technical imperfection of the obesity indices as measures of adiposity, but also to the biologic correlation of lean body mass and percent body fat. Thus, the possibility must be considered that an observed association between any measure of obesity and disease is due to an association with lean body mass rather than adiposity.

Bioelectric Resistance and Conductance

In recent years great interest has developed in the use of bioelectric resistance and conductance measurements to estimate lean body mass and body fat composition (i.e., percent body fat). These measurements are based on the principle that the lean body mass, which consists leargely of ions in a water solution, conducts electricity far better than does fat tissue (van Itallie et al., 1986). Therefore, the resistance of the body to an electrical current is inversely related to the lean body mass. Such measurements should, therefore, provide the same information as obtained with deuterium oxide or other dilution methods. If the total body mass is known, the fat mass and percent body fat can easily be calculated. Electrical resistance is affected also by body shape, so that correcting measurements for height using empirically derived regression formulas or the ratio height^2/resistance can improve the prediction of body composition.

The bioelectrical resistance (or impedance) method is extremely simple in practice. Electrodes (either two or four) are attached to a person's extremities while recumbent but clothed. A weak radio frequency signal is applied to the electrodes and the resistance is measured. Usually several measurements are made, but the whole procedure takes less than 1 minute. The signal generator and recording device (which costs a few thousand dollars) is the size of an attache case and is thus fully portable.

Another method, the "total body electrical conductivity" technique (van Itallie et al., 1986; Segal et al., 1985) is also based on the differential electrical properties of lean and fat mass. The measuring device consists of a large coil

into which a radio frequency current is injected. When a conducting material, such as a body, is passed though the coil, the impedance to the radio frequency current is decreased in relation to the mass of conducting material. Because a person is simply passed through a large coil, no electrodes are required and the procedure is very rapid. The device is, however, expensive and not portable.

Because the resistance and conductance measurements are extremely simple, rapid, safe, and (for the resistance method) portable, they are potentially useful in epidemiologic settings. Their reproducibility and validity, therefore, deserve considerable attention. As discussed for combinations of weight and height, validity in the epidemiologic context is most realistically evaluated by examining the association of the new method with a gold standard after controlling for age and sex (as these are controlled for in any epidemiologic analysis) and in a population with a typical distribution of body fat composition (as opposed to a population with artifically inflated heterogeneity in body composition). Validation studies have generally used densitometry as the standard method, although dilution methods have also been employed.

Measurements of bioelectrical resistance appear to be highly reproducible. Among a population of overweight men and women, Helenius and colleagues (1987) observed an extremely high correlation ($r = 0.99$) for repeated measurements made on the same day. Even in a more homogenous group of 37 healthy young men measured at an interval of 5 days, Lukaski and colleagues (1984) also observed a correlation of 0.99 for replicate measures. Although resistance measurements appear to be remarkably reproducible, their validity as a measure of absolute and relative body components is a more complex issue.

Among 37 healthy young men, Lukaski and colleagues (1985) found resistance measurements (expressed as height^2/resistance) to be highly correlated with total body water ($r = 0.95$). Kushner and Schoeller (1986) found similarly high correlations between height^2/resistance and total body water measured with deuterium ($r = 0.96$ for men and $r = 0.85$ for women). Lukaski and co-workers (1986) compared measurments of both lean body mass and percent body fat based on resistance with the same parameters based on densitometry. Among 84 men, the correlation was 0.98 for fat-free mass and 0.93 for percent body fat. Among 67 women the correlation was 0.95 for fat-free mass and 0.88 for percent body fat. In this study, the estimation of percent body fat among men based on resistance ($r = 0.93$) was only slightly better than that based on four skinfolds (0.88), although this difference was statistically significant. Segal and colleagues (1985) reported correlations between height^2/resistance and densitometrically determined lean body mass of 0.82 for men and 0.92 for women. Based on a much larger sample, Hodgdon and Fitzgerald (1987) observed correlations between percent body fat based on resistance and densitometry measurements of 0.79 among 403 men and 0.82 among 135 women.

In contrast with these promising reports, Helenius and colleagues (1987) reported that, among a population of overweight, middle-aged subjects, resistance (and height^2/resistance) did not significantly predict percent body fat in regression analyses. These authors attributed the differences in findings to the fact that most other investigators had studied healthy, young adults. Their analyses, however, appear to be conceptually inappropriate as resistance is a measure

of lean body mass, not percent body fat. Thus, when attempting to predict percent body fat in a multiple regression analysis, the resistance measurement should only be entered after weight has been forced into the statistical model. A reanalysis of these data would be worthwhile.

At this time, none of the validation studies of resistance measurements have been analyzed in a manner that makes their epidemiologic relevance directly interpretable. In general, correlations with lean body mass have tended to be relatively strong, and higher than for percent body fat, presumably because lean body mass has varied more between subjects in these studies than has percent body fat. The correlations observed in the largest study (Hodgdon and Fitzgerald, 1987) are similar to previously published data unadjusted for age relating weight/height^2 to percent body fat determined by densitometry (see Table 10–3). Thus, electrical resistance measurements do not appear to be sufficiently accurate to replace densitometry as a gold standard and it is not yet clear whether they can provide a substantially better estimate of percent body fat in epidemiologic settings than simple combinations of weight and height.

Data relating to the validity of the "total body conductance" method are far fewer, which is probably related to the limited availability of the instrument. Segal and colleagues (1985) found the correlation between lean body mass determined by densitometry and by conductivity to be 0.93 among 34 men and 0.95 among 41 women. These correlations were higher than those observed among the same subjects using height^2/resistance (0.82 for men and 0.92 for women). Among the men and women combined, the correlation between percent body fat measured by densitometry and by electrical conductivity was 0.97; the correlation between the same densitometry measurement and body composition estimated by electrical resistance was 0.93. Although these results are promising, it must be noted that combining sex and age groups (17 to 59 years) created an unusually large variation in the percentage of body fat (SD = 14.2%), which would tend to inflate correlation coefficients.

Segal and colleagues (1985) also computed the standard error of the electrical conductivity method ($s = \sqrt{\Sigma(Y_o - Y_e)^2/N-2}$, where Y_o is the observed conductivity measurement and Y_e is the conductivity measurement predicted by densitometry; this has been referred to elsewhere as the standard deviation of the residual, see Chapter 6); this standard error was found to be 3.73 percent body fat. Because it has been estimated that the standard error of the densitometry method compared with theoretical truth is on the order of 2.5 percent body fat (Lohman, 1981), Segal and co-workers point out that an appreciable part of the apparent error of the electrical conductivity method may be due to error in the densitometry gold standard. If correct, the validity of all methods that have been compared with densitometry as the standard has been underestimated.

Given the limited data addressing the validity of the resistance and electrical conductivity methods of measuring body composition, and the existence of some uncertainty in the error of the standard method, the role of these measurements in epidemiologic studies remains unclear. Because they are easier and more rapid than standard methods, such as densitometry and deuterium dilution, their validity deserves examination in greater detail from an epidemiologic perspective.

DISTRIBUTION OF BODY FAT

In recent years, considerable attention has been given to the possibility that the distribution of fat is related to certain diseases independent of overall obesity. For example, abdominal fat is independently associated with risk of diabetes (Hartz et al., 1984). Similarly, Vague (1956) has reported that women having a high ratio of skinfold thickness of the posterior neck to thickness at the sacrum (which was termed android obesity) had a high probability of diabetes, gout, and atherosclerosis compared with women with a low ratio (termed gynecoid obesity). The approaches for measuring the distribution of body fat are many as one can imagine a large number of ways in which the distribution of fat might vary, and these distributions may, in turn, be measured in different ways. For example, fat distribution can vary with respect to the proportions stored subcutaneously, intraabdominally, and within the muscle mass. Moreover, the subcutaneous distribution of fat itself can vary between persons; the relative thickness may differ between extremity to trunk and between upper extremity to lower extremity. Understandably, the many possibilities for measurement have led to a plethora of approaches that cannot be reviewed here; however, two of the most commonly employed measures, the waist-to-hip circumference ratio and the biceps-to-subscapular skinfold ratio are discussed.

Waist-to-Hip Circumference Ratio

Individuals with a predominance of abdominal fat exhibit numerous metabolic differences, including insulin resistance and elevated free fatty acid production, compared with those having fat primarily distributed subcutaneously over the lower extremities (Bjorntorp, 1987). These metabolic differences provide a conceptual rationale for examining diseases in relation to these different adipose distributions (Lapides et al., 1984). The most commonly employed measure has been the waist-to-hip circumference ratio.

A standard protocol for measurement of waist-to-hip ratio has not been developed, in part because measurements have been made by personnel with different levels of training, or by subjects themselves in studies conducted by mail. Trained personnel can use bony landmarks for measurements (e.g., halfway between the lower rib margin and the iliac crest), whereas less trained observers may simply be asked to measure the waist at its narrowest point. The specific location of these measurements does appear to provide different results as the correlation of waist-to-hip circumference ratios measured with different landmarks was 0.77 for women and 0.90 for men (Seidell et al., 1987). Factors, such as postprandial status, time of day, standing position, and depth of inspiration, also affect these measurements. Because the degree to which these factors may contribute error is uncertain, standardizing such factors within a study to the greatest degree possible is prudent. The investigation of sources of error in this measurement deserves further attention.

The validity of the waist-to-hip circumference ratio as a measure of abdom-

inal obesity was evaluated in two studies by comparing this ratio with abdominal fat measured by computed tomography (Ashwell et al., 1985; Seidell et al., 1987). Reasonably strong correlations were seen in both; in the larger study (Seidell et al., 1987), the correlation coefficients were 0.55 for women and 0.75 for men. Age alone, however, predicted intraabdominal fat nearly as well as the waist-to-hip ratio, so that after adjustment for age and body mass index, the waist-to-hip ratio was not significantly correlated with intraabdominal fat. Because the focus of epidemiologic studies is usually the effect of abdominal fat independent of age and overall obesity, this analysis suggests that the errors associated with the use of waist-to-hip ratios in epidemiologic studies may be large. In a study among 28 women, Ashwell and colleagues (1985) observed a correlation of 0.69 between the waist-to-hip circumference ratio and intraabdominal fat measured cross-sectionally by computed tomography. This remained statistically significant, although diminished, after adjustment for age and overall obesity. The inconsistency of these studies indicates the need for further assessment of the validity of the waist-to-hip circumference ratio using computed tomography as a standard. If used as an index of intraabdominal fat, consideration should be given to the possibility that the ratio of waist and hip circumference is not the optimal combination of these variables. For example, waist circumference, alone or adjusted for height, may possibly be superior.

It is notable that in a prospective study where the association of waist-to-hip circumference ratios with cardiovascular disease was both statistically significant and important (relative risks between two and three), the difference in means between subjects who subsequently developed disease and those who remained healthy (0.938 vs. 0.925) was only about 1 percent (Larsson et al., 1984). (This can happen when the distribution of exposure is narrow, i.e., the standard deviation is small in relation to the mean.) Because even the smallest degree of systematic bias or an effect of disease on these measurements could obscure such a difference, it is uncertain whether valid case-control studies can be conducted using this ratio (see Chapter 2).

Biceps and Subscapular Skinfolds

The triceps and subscapular skinfold thicknesses have been used as relative measures of extremity and truncal obesity, respectively. In a number of studies it has been suggested that truncal obesity is more strongly related to disorders of carbohydrate and lipid metabolism, and hypertension than is peripheral obesity (Blair et al., 1984). The use of triceps and subscapular skinfolds appears to be largely based on convention and convenience; it is possible that skin folds measured at the other sites might be more representative of obesity in the extremities or trunk and thus, stronger predictors of disease (Roche, 1984).

The distinction between intraabdominal and subcutaneous truncal adiposity with respect to disease occurrence is not well-defined at present. The interpretation of data relating to body fat distribution is difficult as the measurements at different areas tend to be strongly correlated. It is possible, for example, that subscapular skin folds are simply acting as a surrogate for intraabdominal fat.

To understand the physiologic effects of fat distribution, it is important to use a number of these measures simultaneously. The interpretation of such studies is likely to be further complicated as the strength of associations between various measures of fat distribution depend on the relative degree of measurement error associated with each of them, as well as the true strength of their biologic relationships.

SUMMARY

Most of our present information on the health effects of obesity is derived from data on weight and height. Because these variables are easy and inexpensive to collect and can be assessed quite accurately even by self-report, they continue to play a central role in epidemiologic studies. Height and other dimensions that potentially reflect early nutritional status could be exploited more fully than has been done thus far. When using combinations of weight and height to assess obesity, Quetelet's index (weight/height^2) is as strongly correlated with percent body fat as alternative formulations and provides the considerable advantage of comparability across studies.

Despite the widespread use of obesity indices based on weight and height, their error in representing percent body fat is considerable, not because of failure to measure height and weight accurately, but because the lean body mass can vary considerably among persons of the same height. Indeed, the validity of combinations of weight and height for epidemiologic applications, assessed by correlations with densitometry, has probably been overestimated in the past by artificially increasing the between-person variation in fatness by using study populations heterogeneous with respect to age and sex. Within groups of similar age and sex, the correlation between Quetelet's index and densitometry appears to be approximately 0.60 for men, but possibly somewhat higher for women. However, the assumption that percent body fat is more physiologically relevant than absolute fat mass (adjusted for a measure of body size) deserves to be questioned.

The degree of error associated with obesity indices implies that the health effects of obesity may have been substantially underestimated in epidemiologic studies. This underestimation can be addressed by statistical corrections for measurement error, as discussed in Chapter 12, or by using improved methods to measure percent body fat in future studies. Among the alternative approaches, bioelectric impedance appears to be the only method sufficiently simple and inexpensive to be widely used epidemiologically. Sufficient data, however, do not exist to quantify the incremental accuracy it can provide among populations typically studied by epidemiologists.

The present interest in distribution of body fat is likely to enhance our understanding of the pathophysiology of numerous diseases. Much additional work, however, is needed to define what is actually being measured by the ratios that are commonly employed today, as they may reflect both total fatness as well as fatness in specific locations. The public health implications of body fat distribution are unclear as fat reduction interventions are not focused at specific sites.

REFERENCES

Allen, T. H., M. T. Peng, K. P. Chen, T. F. Huang, C. Chang, and H. S. Fang (1956). Prediction of blood volume and adiposity in man from body weight and cube of height. *Metabolism 5,* 328–345.

Ashwell, M., T. J. Cole, and A. K. Dixon (1985). Obesity: New insight into the anthropometric classification of fat distribution shown by computed tomography. *Br. Med. J. 290,* 1692–1694.

Benn, R. T. (1971). Some mathematical properties of weight-for-height indices used as measures of adiposity. *Br. J. Prev. Soc. Med. 25,* 42–50.

Bjorntorp, P. (1987). Classification of obese patients and complications related to the distribution of surplus fat. *Am. J. Clin. Nutr. 45 (s),* 1120–1125.

Blair, D., J. P. Habicht, E.A.H. Simms, D. Sylwester, and S. Abraham (1984). Evidence for an increased risk for hypertension with centrally located body fat and the effect of race and sex on this risk. *Am. J. Epidemiol. 119,* 526–540.

Bray, G. A., F. L. Greenway, M. E. Molitch, W. T. Dahms, R. L. Atkinson, and K. Hamilton (1978). Use of anthropometric measures to assess weight loss. *Am. J. Clin. Nutr. 31,* 769–773.

Brinkley, D., R. G. Carpenter, and J. L. Haybittle (1971). An anthropometric study of women with cancer. *Br. J. Prev. Soc. Med. 25,* 67–75.

Brozek, J., F. Grande, J. T. Anderson, and A. Keys (1963). Densitometric analysis of body composition: Revision of some quantitative assumptions. *Ann. N.Y. Acad. Sci. 110,* 113–140.

Colliver, J. A., S. Frank, and A. Frank (1983). Similarity of obesity indices in clinical studies of obese adults: A factor analytic study. *Am. J. Clin. Nutr. 38,* 640–647.

Criqui, M., H. Klauber, E. Barrett-Connor, M. J. Holdbrook, L. Suarez, and D. L. Wingard (1982). Adjustment for obesity in studies of cardiovascular disease. *Am. J. Epidemiol. 116,* 685–691.

de Waard, F. (1975). Breast cancer incidence and nutritional status with particular reference to body weight and height. *Cancer. Res. 35,* 3351–3356.

Evans, J. G. and I.A.M. Prior (1969). Indices of obesity derived from height and weight in two Polynesian populations. *Br. J. Prev. Soc. Med. 23,* 56–59.

Fanelli, M. T. and R. J. Kuczmarski (1984). Ultrasound as an approach to assessing body composition. *Am. J. Clin. Nutr. 39,* 703–709.

Florey, C. du V. (1970). The use and interpretation of ponderal index and other weight–height ratios in epidemiological studies. *J. Chron. Dis. 23,* 93–103.

Frisancho, A. R. (1984). New standards of weight and body composition by frame size and height for assessment of nutritional status of adults and the elderly. *Am. J. Clin. Nutr. 40,* 808–819.

Frisch, R. E. and W. J. McArthur (1974). Menstrual cycles: Fatness as a determinant of minimum weight for height necessary for their maintenance or onset. *Science 185,* 949–951.

Garn, S. M. Optimal Nutritional Assessment. (1979) In D. B. Jellifee and E.F.P. Jellifee, eds.: *Human Nutrition, A Comprehensive Treatise: Nutrition and Growth.* New York: Plenum Press pp. 273–296.

Garn, S. M. and S. D. Pesick (1982). Comparison of the Benn index and other body mass indices in nutritional assessment. *Am. J. Clin. Nutr. 36,* 573–575.

Garn, S. M., W. R. Leonard, and V. M. Hawthorne (1986). Three limitations of the body mass index. *Am. J. Clin. Nutr. 44,* 996–997.

Garrow, J. S. (1983). Indices of adiposity. *Nut. Abst. Rev. 53,* 697–708.

Goldbourt, U. and J. H. Medalie (1974). Weight–height indices. Choice of the most suitable index and its association w/selected variables among 10,000 adult males of heterogeneous origin. *Br. J. Prev. Soc. Med. 28,* 116–126.

Grodin, J. M., P. K. Siiteri, and P. C. MacDonald (1973). Source of estrogen production in postmenopausal women. *J. Clin. Endocrinol. Metab. 36,* 207–214.

Habicht, J. P., C. Yarbrough, and R. Martorell (1979). Anthropometric field methods: Criteria for selection. In D. G. Jelliffe and E.F.P. Jelliffe, eds.: Nutrition and Growth. *Human Nutrition 2,* 365–387.

Harsha, D. W. and R. R. Frerichs, G. S. Berenson (1978). Densitometry and anthropometry of black and white children. *Human Biology 5,* 261–280.

Hartz, A. J., D. C. Rupley, and A. A. Rimm (1984). The association of girth measurements with disease in 32,856 women. *Am. J. Epidemiol. 119,* 71–80.

Helenius, M.Y.T., D. Albanes, M. S. Micozzi, P. R. Taylor, and O.P.F. Heinonen (1987). Studies of bioelectric resistance in overweight, middleaged subjects. *Human Biology 59,* 271–279.

Hirayama, T. (1978). Epidemiology of breast cancer with special reference to the role of diet. *Prev. Med. 7,* 173–195.

Hodgdon, J. A., P. I. Fitzgerald (1987). Validity of Impedance Predictions at various levels of fatness. *Hum Bio 59,* 281–298.

Insull, W., Jr., T. Oiso, and K. Tsuchiya (1968). Diet and nutritional status of Japanese. *Am. J. Clin. Nutr. 21,* 753–777.

Katch, V. L. and P. S. Freedson (1982). Body size and shape: Derivation of the "HAT" frame size model. *Am. J. Clin. Nutr. 36,* 669–675.

Katch, V. L., P. S. Freedson, F. I. Katch, and L. Smith (1982). Body frame size: Validity of self appraisal. *Am. J. Clin. Nutr. 36,* 676–679.

Keys, A., F. Fidanza, M. J. Karvonen, N. Kimura, and H. L. Taylor (1972). Indices of relative weight and obesity. *J. Chron. Dis. 25,* 329–343.

Khosla, T. and C. R. Lowe (1967). Indices of obesity derived from body weight and height. *Br. J. Prev. Soc. Med. 21,* 122–128.

Killeen, J., D. Vanderburg, and W. Harlan (1978). Application of weight-height ratios and body indices to juvenile populations—The National Health Examinations Survey Data. *J. Chron. Dis. 31,* 529–537.

Kuczmarski, R. J., M. T. Fanelli, and G. G. Koch (1987). Ultrasonic assessment of body composition in obese adults: Overcoming the limitations of the skinfold caliper. *Am. J. Clin. Nutr. 45,* 717–724.

Kushner, R. F. and D. A. Schoeller (1986). Estimation of total body water by bioelectrical impedance analysis. *Am. J. Clin. Nutr. 44,* 417–424.

Lapidus, L., C. Bengtsson, B. Larsson, K. Pennert, E. Rybo, and L. Sjostrom (1984). Distribution of adipose tissue and risk of cardiovascular disease and death: A 12-year follow-up of participants in the population study of women in Gothenburg, Sweden. *Br. Med. J. 289,* 1257–1261.

Larsson, B., K. Svardsudd, L. Welin, L. Wilhelmsen, P. Bjorntorp, and G. Tibblin (1984). Abdominal adipose tissue distribution, obesity, and risk of cardiovascular disease and death: 13-year follow-up of participants in the study of men born in 1913. *Br. Med. J. 288,* 1401–1404.

Lee, J., L. N. Kolonel, and M. W. Hinds (1981). Relative merits of the weight-corrected-for-height indices. *Am. J. Clin. Nutr. 34,* 2521–2529.

Lew, E. A. and L. Garfinkel (1979). Variations in mortality by weight among 750,000 men and women. *J. Chron. Dis. 32,* 563–576.

Lohman, T. G. (1981). Skinfolds and body density and their relation to body fatness: A review. *Hum. Biol. 53,* 181–225.

Lohman, T. G., A. F. Roche, and R. Martorell (1988). *Anthropometric Standardization Reference Manual.* Champaign, Il.: Human Kinetics Books.

Lukaski, H. C. and P. E. Johnson (1985). A simple, inexpensive method of determining total body water using a tracer dose of D_2O and infra-red absorption on biological fluids. *Am. J. Clin. Nutr. 41,* 363–370.

Lukaski, H. C., P. E. Johnson, W. W. Bolonchuk, and G. I. Lykken (1985). Assessment of fat-free mass using bioelectrical impedance measurements of the human body. *Am. J. Clin. Nutr. 41,* 810–817.

Lukaski, H. C., W. W. Bolonchuk, C. B. Hall, and W. A. Siders (1986). Validation of tetrapolar bioelectrical impedance method to assess human body composition. *J. Appl. Physiol 60,* 1327–1332.

Lukaski, H. C. (1987). Methods for the assessment of human body composition: Traditional and new. *Am. J. Clin. Nutr. 46,* 537–556.

Manson, J. E., M. J. Stampfer, C. H. Hennekens, and W. C. Willett (1987). Body weight and longevity: A reassessment. *J.A.M.A. 257,* 353–358.

Mellits, E. D. and D. B. Cheek (1970). The assessment of body water and fatness from infancy to adulthood. In J. Brozek, ed.: Monographs of the Society for Research in Child Development, Vol. 35, No. 7. pp. 12–26.

Metropolitan Life Insurance Co. (1959). New weight standards for men and women. *Stat Bull N.Y. Metropolitan Life Insurance Co.* 40:1–4.

Michielutte, R., R. A. Diseker, W. T. Corbett, H. M. Schey, and J. R. Ureda (1984). The relationship between weight–height indices and the triceps skinfold measure among children age 5 to 12. *Am. J. Public Health,* 74, 604–606.

Micozzi, M. S. (1985). Nutrition, body size, and breast cancer. *Yearbook of Physical Anthropology 28,* 175–206.

Micozzi, M. S., D. Albanes, D. Y. Jones, and W. C. Chumlea (1986). Correlations of body mass indices with weight, stature, and body composition in men and women in NHANES I and II. *Am. J. Clin. Nutr. 44,* 725–731.

Moore, F. D., K. H. Olesen, J. D. McMurrey, H. V. Parker, M. R. Ball, and C. M. Boyden (1963). *The Body Cell Mass and Its Supporting Environment.* Philadelphia: W. B. Saunders.

Mueller, W. H. and L. Stallones (1981). Anatomical distribution of subcutaneous fat: skinfold site choice and construction of indices. *Human Biology 53,* 321–335.

Parizkova, J. (1961). Total body fat and skinfold thickness in children. *Metabolism 10,* 794–807.

Palta, M., R. J. Prineas, R. Berman, and P. Hannan (1982). Comparison of self-reported and measured height and weight. *Am. J. Epidemiol. 115,* 223–230.

Quetelet, L.A.J. (1869). *Physique Sociale.* Brussels: C. Muquardt, Vol 2, p. 92.

Revicki, D. A. and R. G. Israel (1986). Relationship between body mass indices and measures of body adiposity. *Am. J. Public Health 76,* 992–994.

Rhoads, G. G. and A. Kagan (1983). The relation of coronary disease, stroke, and mortality to weight in youth and middle age. *Lancet 1,* 492–495.

Roche, A. F., R. M. Sievogel, W. Chumlea, and P. Webb (1981). Grading body fatness from limited anthropometric data. *Am. J. Clin. Nutr. 34,* 2831–2838.

Roche, A. F. (1984). Anthropometric methods: new and old, what they tell us. *Int. J. Obesity 8,* 509–523.

Roche, A. F., R. N. Baumgartner, and S. Guo (1987). Population methods: Anthropometry or estimations. In N. G. Norgan, ed.: *Human Body Composition and Fat Distri-*

bution. Report of an EC Workshop, London 10–12 December, 1985. CIP-gegevens Koninklijke Bibliotheek, Den Haag, The Netherlands.

Rose, G. A., H. Blackburn, R. F. Gillum, and R. J. Prineas (1982). Cardiovascular survey methods. WHO Monograph Series No. 56.

Ruiz, L., J. R. Colley, and P. J. Hamilton (1971). Measurement of triceps skinfold thickness. An investigation of sources of variation. *Brit. J. Prev. Soc. Med. 25,* 165–167.

Segal, K. R., B. Gutin, E. Presta, J. Wang, and T. B. van Itallie (1985). Estimation of human body fat composition by electrical impedance methods: A comparative study. *J. Appl. Physiol. 58,* 1565–1571.

Seidell, J. C., A. Oosterlee, M.A.O. Thijssen, J. Burema, P. Deurenberg, J. G. Hautvast, and J.H.J. Ruijs (1987). Assessment of intra-abdominal and subcutaneous abdominal fat: Relation between anthropometry and computed tomography. *Am. J. Clin. Nutr. 45,* 7–13.

Seltzer, C. C., R. F. Goldman, and J. Mayer (1965). The triceps skinfold as a predictive measure of body density and body fat in obese adolescent girls. *Pediatrics 36,* 212–218.

Sheng, H. P. and R. A. Huggins (1979). A review of body composition studies with emphasis on total body water and fat. *Am. J. Clin. Nutr. 32,* 630–647.

Siri, W. E. (1961). Body composition from fluid spaces and density: Analysis of methods. In Techniques For Measuring Body Composition. Nat'l. Acad. Sci., Washington, D.C.: National Research Council, pp. 223–244.

Stewart, A. L. (1982). The reliability and validity of self-reported weight and height. *J. Chron. Dis. 35,* 295–309.

Stewart, A. W., R. T. Jackson, M. A. Ford, and R. Beaglehole (1987). Underestimation of relative weight by use of self-reported height and weight. *Am. J. Epidemiol. 125,* 122–126.

Stunkard, A. J. and J. M. Albaum (1981). The accuracy of self-reported weight. *Am. J. Clin. Nutr. 34,* 1593–1599.

Vague, J. (1956). The degree of masculine differentiation of obesities: A factor determining predisposition to diabetes, atherosclerosis, gout and uric calculous disease. *Am. J. Clin. Nutr. 4,* 20–34.

van Itallie, T. B., K. R. Segal, and R. C. Funk (1986). Total body electrical conductivity: A rapidly measured index of lean body mass. In N. G. Norgan, ed.: *Human Body Composition and Fat Distribution.* Euro-Nutr. Report No. 8 (Wageningen, Holland: Euro-Nutr.), pp. 113–127.

Watson, P. E., I. D. Watson, and R. D. Batt (1980). Total body water volumes for adult males and females estimated from simple anthropometric measurements. *Am. J. Clin. Nutr. 33,* 27–39.

Womersley, J. (1977). A comparison of the skinfold method with extent of "overweight" and various weight–height relationships in the assessment of obesity. *Br. J. Nutr. 38,* 271–284.

Zerfar A. (1979). Anthropmetric field methods: General. In D. B. Jellife and E.F.P. Jellife, eds.: *Human Nutrition, A Comprehensive Treatise: Nutrition and Growth.* New York: Plenum Press, pp. 339–362.

11

Implications of Total Energy Intake for Epidemiologic Analyses

Total energy intake deserves special consideration in nutritional epidemiology for three reasons:

1. The level of energy intake may be a primary determinant of disease.
2. Individual differences in total energy intake produce variation in intake of specific nutrients unrelated to dietary composition as the consumption of most nutrients is positively correlated with total energy intake. This added variation may be extraneous, and thus a source of error, in many analyses.
3. When energy intake is associated with disease but is not a direct cause, the effects of specific nutrients may be distorted, that is, confounded, by total energy intake.

Before examining these three issues in detail, the physiologic aspects of energy utilization and the determinants of variation in energy intake in epidemiologic studies are discussed. In accordance with common practice, total caloric intake is used synonymously with total energy intake in this chapter, although use of joules as a unit of measurement is more correct and most journals now require SI units.

PHYSIOLOGIC DETERMINANTS OF ENERGY UTILIZATION

Physiologists have partitioned energy expenditure into several components: resting metabolic rate, thermogenic effect of food, physical activity, and adaptive thermogenesis (Horton, 1983) (Fig. 11–1). Resting metabolic requirements are

Parts of this chapter are reproduced from Willett and Stampfer (1986) and Willett (1987) with permission of the original publishers. Several of the ideas on various analytic strategies evolved during a March, 1988, workshop sponsored by the SEARCH program of the International Agency for Cancer Research, Lyon, France.

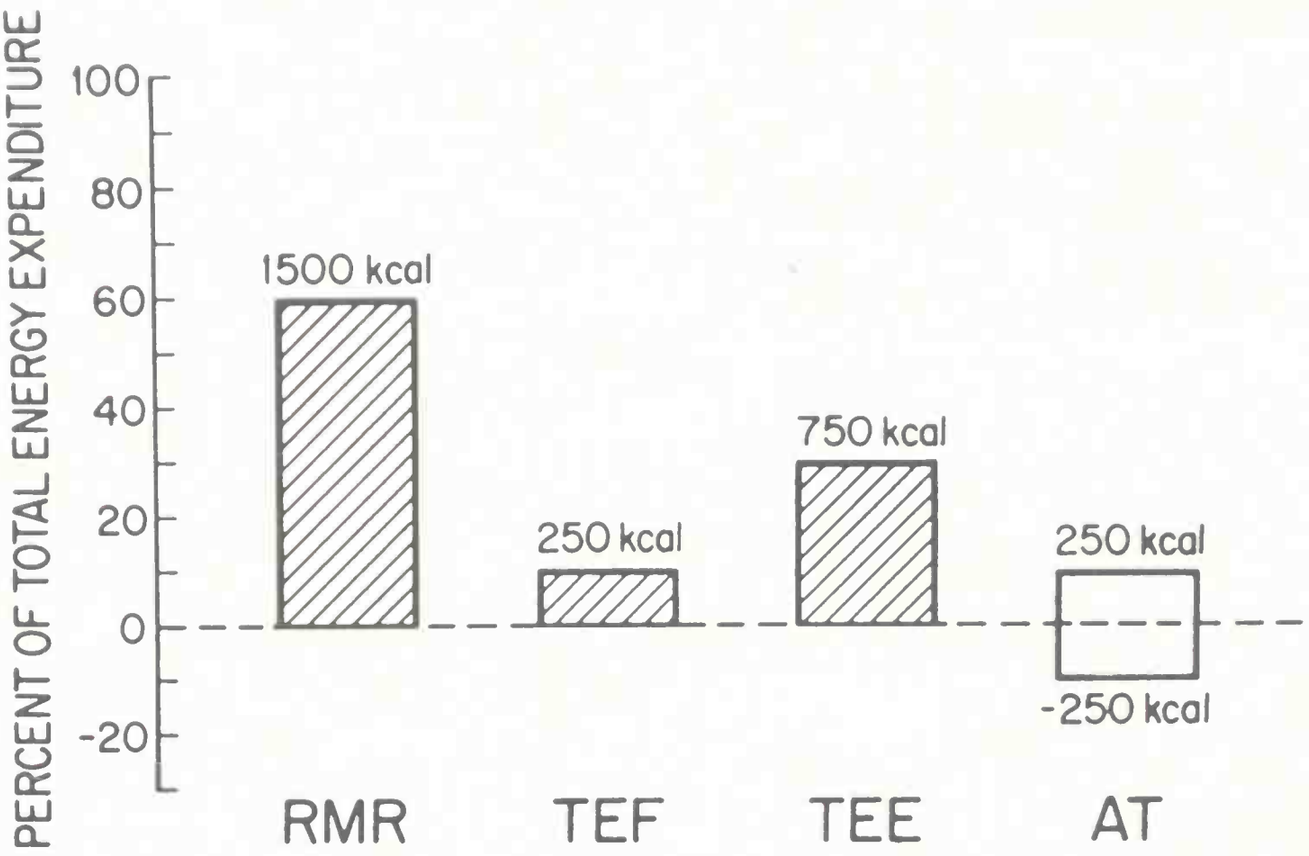


Figure 11–1. Components of energy expenditure. (From Horton, 1983, reproduced with permission.)

quantitatively the most important, representing approximately 60 percent of total energy expenditure in most individuals. The thermogenic effect of food (which is the metabolic cost of absorbing and processing carbohydrate, protein, and fat) varies with the sources of energy (Donato and Hegsted, 1985) but is only about 10 percent of the total. Adaptive thermogenesis represents the capacity of an individual to conserve or expend energy in response to variable intake of food or, perhaps, temperature extremes. In humans, adaptive thermogenesis is defined differently by various investigators (Sjostrom, 1985) and is difficult to measure, but it has been estimated to represent a maximum of ± 10 percent of calories. In a moderately active individual, physical activity accounts for approximately 30 percent of energy intake (Horton, 1983).

Determinants of Between-Person Variation in Total Energy Intake

Although it is helpful to consider the average values for physiologic components of energy expenditure, epidemiologists are primarily interested in the determinants of *variation* in energy intake between individuals. Although many specific factors influence energy intake, they can be considered as three general categories: body size, metabolic efficiency, and physical activity. Departures from energy balance, that is, change in body energy stores due to intake above or below expenditure, also account for part of the observed variation among persons (Fig. 11–2).

Body size affects the amount of energy needed for resting metabolic activity as well as to sustain physical exertion. On the basis of careful measurements by Jequier and Schutz (1983), using a specially designed respiratory chamber, it

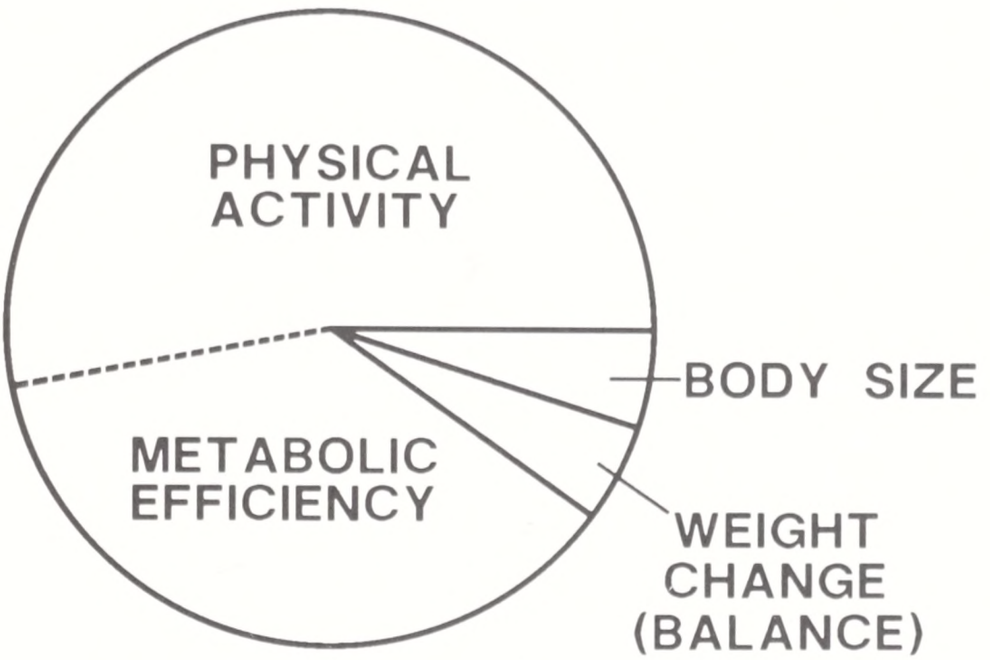


Figure 11–2. Components of between-person variation in energy intake. The relative size of these components varies depending on the population being studied.

appears that body size is a primary determinant of energy expenditure (Fig. 11–3), particularly at low levels of physical activity. These authors found that 24-hour energy expenditure was a linear function of body weight, which accounted for 74 percent of the variance in expenditure. Using similar methods, it has further been demonstrated that energy expenditure is primarily related to lean body

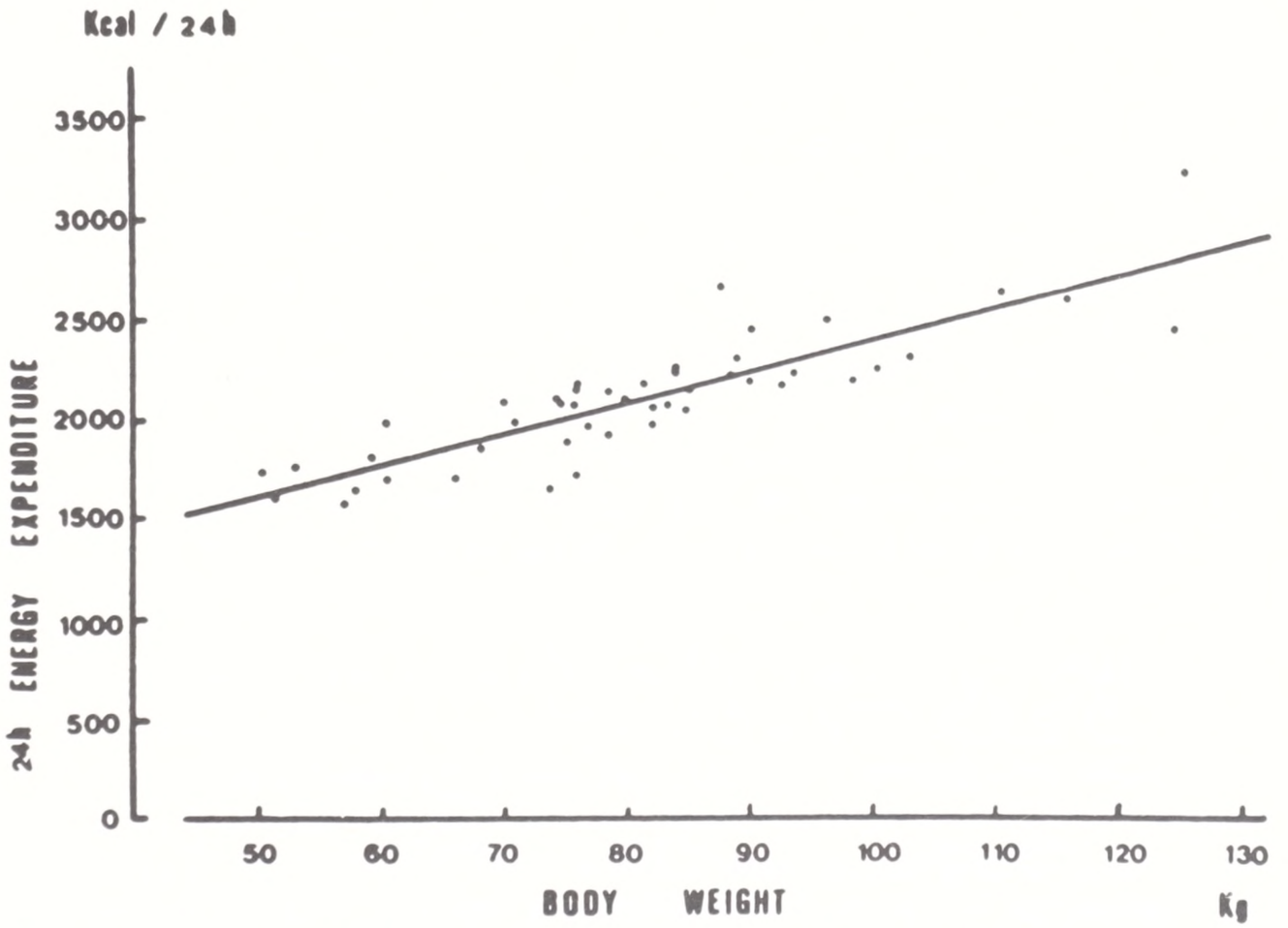


Figure 11–3. Relation between body weight and 24-hour energy expenditure, based on respiration-chamber measurements. (From Jequier and Schultz, 1983; reproduced with permission.)

mass rather than to fat mass (Ravussin et al., 1986). It must be noted that these study groups were heterogeneous with respect to age and gender and were atypically enriched with obese subjects, all of which would tend to inflate variation attributable to weight or lean body mass. More important, as noted by these authors, the usual physical activity of subjects in these studies was constrained by their restriction to a small metabolic chamber.

Among free-living subjects neither height nor weight accounts for a major proportion of the between-person variation in total caloric intake (Thompson and Billewicz, 1961; Gordon et al., 1984). For example, neither measure of body size was significantly correlated with caloric intake among a group of 194 women aged 34 to 59 years who collected four 1-week weighed diet records over 1 year (Willett et al., 1985) (Spearman $r = 0.08$ for height and $r = -0.10$ for weight, unpublished data). Underreporting of intake by clearly obese subjects has been documented (Prentice et al., 1986) and could explain some of the lack of a positive association between weight and energy intake. Body size, however, does not appear to be the major determinant of variation in energy intake among free-living subjects within a specific age and sex group. (If analyses are conducted on groups that are markedly heterogeneous in size, which would be true if both men and women were included, then body size would become a more important determinant of variation in energy expenditure.)

Although energy expenditure at rest accounts for the majority of absolute energy intake, it does not vary greatly among individuals of similar age and sex. Thus, physical activity assumes a relatively large role in determining the variation in energy expenditure among individuals unconstrained by a metabolic chamber. The positive relation between physical activity and energy intake has been appreciated for years (Johnson et al., 1956) and has been clearly demonstrated by Morris and colleagues (1977) in an epidemiologic study among a population of bank clerks (Table 11–1). Among the 194 women previously mentioned, the correlation between a physical activity questionnaire and caloric intake based on 28 days of diet recording was 0.22 (unpublished data). Because the obese tend to be less physically active, this relationship is sufficient to explain the weak inverse relation between relative weight and caloric intake that has been observed in most epidemiologic studies (Gordon et al., 1981; Romieu et al., 1988). The relationship also presumably underlies the decline in per capita caloric intake in the United States despite an increasing prevalence of obesity (Van Itallie, 1978; National Center for Health Statistics, 1979; Abraham and

Table 11–1. Association of leisure time physical activity and energy intake among London bus drivers

Thirds of distribution of men for energy intake	Physical activity in leisure time		
	Little	Moderate	Much
Low third	16	9	4
Middle third	9	12	3
High third	5	10	9

From Morris et al., 1977.

Carroll 1979). The true proportion of variation in energy intake accounted for by physical activity varies substantially among populations and is likely to be seriously underestimated in most studies because of difficulty in accurately measuring physical activity. Recently, Ravussin and colleagues (1986) have demonstrated that even motor activity within the confines of a respiratory chamber ("fidgeting") varies dramatically between persons and can account for hundreds of kilocalories per day. Such differences in activity would not be detected by typical questionnaires. Thus, it appears likely that physical activity, which includes both fine motor as well as major muscle movement, is the dominant explanation for between-person differences in energy intake. Indeed, in most instances total energy intake can be interpreted as a crude measure of physical activity, particularly after controlling for body size, age, and sex.

Metabolic efficiency may contribute to individual differences in caloric intake; metabolically inefficient persons require greater amounts of energy to maintain their level of activity and weight. The mechanisms and determinants of metabolic efficiency (including differences in absorption and the general category of thermogenesis) are poorly defined in humans and are beyond the scope of this chapter. Individual differences apparently exist, however, as under carefully controlled conditions some subjects gain weight more rapidly than others who have a similar caloric intake (Sims et al., 1973). It has been suggested that the obese are more metabolically efficient (James and Trayhurn, 1981); however, evidence is accumulating that their energy requirements are not lower than the nonobese at a similar level of physical activity (Jequier and Schultz, 1983; Garrow and Webster, 1985; Webb, 1985). It is also clear that increases in energy intake reduce metabolic efficiency; in other words, excessive energy intake induces increased thermogenesis. This mechanism thus contributes to the regulation of body weight (Miller, 1973; Himms-Hagen, 1984; Woo et al., 1985). This capacity to alter metabolic efficiency (also called adaptive thermogenesis) is limited and has been estimated to have a range of approximately ± 10 percent of average daily caloric intake (Webb, 1985).

Unfortunately, there is no practical method of measuring metabolic efficiency in an epidemiologic setting. The data of Jequier and Schultz (1983, Fig. 11–3), however, suggest that there is relatively little between-person variation in energy expenditure after physical activity has been restricted and weight is accounted for. Thus, the contribution of metabolic efficiency to between-person variation in energy expenditure remains poorly defined but is not likely to be large.

The net balance of energy intake in relation to body size, metabolic efficiency, and physical activity determines whether a person gains or loses weight. In the absence of compensatory mechanisms, relatively small changes maintained over long periods have a profound effect on body weight. For example, if an adult male who consumes 2500 kcal per day increases his caloric consumption by only 2 percent with other factors remaining constant, over a 10-year period a theoretical 20-kg weight gain will result (theoretical weight change over 10 years = 2500 kcal per day $\times$ 0.02 $\times$ 365 days per year $\times$ 10 years/9,000 kcal per kilogram of fat = 20.3 kg). In reality, the increase in weight will not be so dramatic, as the additional energy cost of maintaining and moving the added

mass eventually equals the increment in energy intake and a new steady-state is obtained.

Hofstetter and colleagues (1986) have used the cross-sectional data in Figure 11–3, in which 1 kg in weight corresponds to about 20 kcal per day, to estimate the ultimate weight gain associated with a given change in energy availability. In this manner, the 2 percent change in energy intake (50 kcal/day) would result in an ultimate change of 2.5 kg, which is still readily measurable epidemiologically. (It must be noted that such calculations of projected weight changes from cross-sectional data are likely to be somewhat inaccurate for several reasons, including that weight in the cross-sectional data is largely lean body mass, whereas changes in weight caused by simple increases in caloric intake would be primarily due to adipose.) Careful, long-term studies of the effects of small increments in energy intake on body weight would be most useful as most studies have been done using large increments (Webb, 1985).

During short periods, such as months, the proportion of energy intake accounted for by balance (i.e., weight gain or loss) is larger if some persons are experiencing rapid weight change. Over the long-term, however, balance can account for only a very small part of between-person differences in energy intake. The true between-person variation in long-term caloric intake is poorly defined. Using 28 days of diet recording per subject (which minimizes variation within persons), we observed a standard deviation of 323 kcal/day (mean 1620) among middle-aged women (Willett et al., 1985). This amount is somewhat less than the standard deviation of 473 kcal/day for women (mean 1793) estimated by Beaton and co-workers (1979) from multiple 24-hour recalls using an analysis-of-variance model.

Because differences in total energy intake between individuals are largely determined by physical activity, body size, and metabolic efficiency, it is apparent that an epidemiologic study of risk of disease simply in relation to energy intake is difficult or impossible to interpret. Energy intake is measurable only crudely (errors being considerably larger than 2%) with standard questionnaires or interviews; physical activity is probably measured even more crudely than diet; and metabolic efficiency is essentially unmeasurable in an epidemiologic setting. Thus, it would be difficult if not impossible to partition an individual's total energy intake with sufficient precision so as to measure the balance available after accounting for physical activity and metabolic efficiency because this balance would be computed as the difference between two crudely measured variables minus an unmeasurable variable. It is not surprising that the degree of obesity is not strongly correlated with energy intake in cross-sectional studies.

Although the interpretation of energy intake data in epidemiologic analyses may be difficult, simple and readily available measures of weight and height can be extremely useful as alternatives to direct measures of energy intake. The presence of high relative weight implies that, sometime in life, a positive balance between energy intake and energy expenditure has occurred. Even more useful, a change in weight implies a positive or negative energy balance during that time. The interpretation of data on weight and height, however, are potentially complicated as individuals apparently have some ability to compensate for increased caloric intake by increasing thermogenesis (i.e., reducing metabolic

efficiency). Moreover, individuals may vary in their capacity to respond in this way (Miller, 1973). It is thus conceivable that excess caloric intake could increase disease risk in a manner that was mediated by a thermogenic response (i.e., related to increased metabolic rate) rather than by the accumulation of fat. The interpretation of epidemiologic data relating to weight and energy intake, therefore, depends in an important way on how body weight responds to increased energy intake. In model 1 (Fig. 11–4), increased energy intake is fully compensated by adaptive thermogenesis up to a certain point, and weight gain occurs only after a threshold increase in caloric intake is exceeded. In model 2, any long-term increase in energy intake causes weight gain; any compensatory increase in thermogenesis occurs only in conjunction with weight gain.

The implication of model 2 for epidemiologists is that the absence of a relationship between obesity and risk of disease implies that the disease is not caused by excess energy intake. Failure to observe an association between relative weight and risk of disease thus provides evidence that changes in energy intake alone do not influence disease occurrence. This may be appreciated by considering the hypothetical situation where even a small increase in energy intake, say 5 percent, with physical activity held constant raises the risk of disease. If model 2 applies, this difference in intake would produce an easily measurable weight gain; therefore, individuals who had increased their energy consumption would weigh more and we would observe a positive association between relative weight and risk of disease. If model 1 were correct, an absence of association with obesity would not exclude the possibility that increased energy intake causes a higher risk of disease by inducing a fully compensating

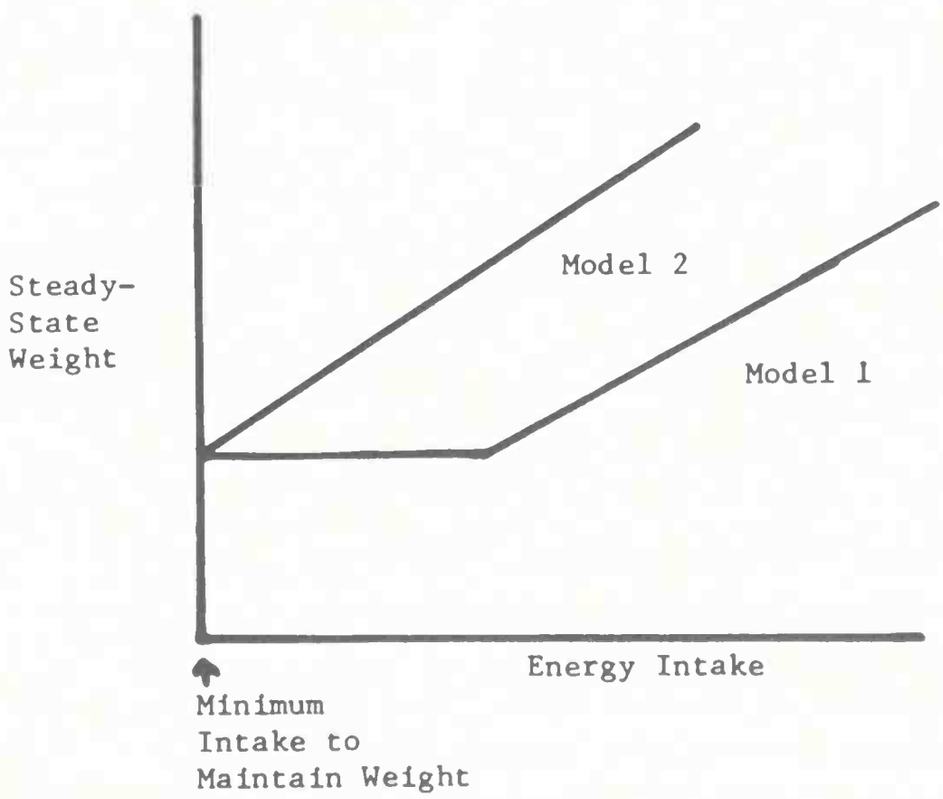


Figure 11–4. Alternate models of response to an increase in long-term energy intake. In model 1, increased thermogenesis prevents weight gain until up to a certain increment in energy intake. In model 2, any compensatory increase in energy expenditure occurs only in conjunction with weight gain.

higher level of thermogenesis. Recognizing that individuals may vary in their response to energy intake, model 2 would need to apply only to an appreciable proportion of the population, not necessarily to all individuals, to be relevant epidemiologically.

The data of Donato and Hegsted (1985), based on rats, indicate that gain in body weight is a linear function of energy availability, that is, model 2 applies. Based on a review of metabolic studies among humans, Woo and colleagues (1983) suggest that any adaptive response in thermogenesis occurs only after some change in adiposity. Although additional human data based on careful metabolic studies would be helpful, it seems appropriate to interpret a lack of association with relative weight in a specific study as evidence against a direct causal effect of total energy intake on risk of disease.

Relation of Energy Intake with Specific Nutrient Intake

Intakes of most nutrients in free-living populations tend to be positively correlated with total energy intake (Jain et al., 1980; Lyons et al., 1983; Gordon et al., 1984). Data based on four 1-week diet records were used to examine these correlations among 194 women (see Table 11–2, last row). Correlations were particularly strong for fat, protein, and carbohydrate (which contribute to energy intake); in this population alcohol intake was quite modest and was only weakly correlated with total energy consumption. Every other nutrient examined, however, was also correlated with total energy intake even though many did not contribute to energy. For example, the correlation with energy was 0.36 for fiber, 0.45 for vitamin A, and 0.34 for vitamin C. This tendency for all nutrients, even minerals and vitamins, to be correlated with total energy intake results from the tendency of larger, more active, and less metabolically efficient persons to eat more food in general.

These interrelations are further complicated by the observation that the composition of diets may vary by level of total caloric intake, depending on the behavior of the population. For example, as shown in Table 11–3, women with lower energy intake tended to have a proportionally higher intake of fiber than women with higher energy intake. These correlations between caloric intake and specific nutrient intakes further highlight the need to consider total energy intake when interpreting associations between specific nutrients and disease in epidemiologic studies.

ADJUSTMENT FOR ENERGY INTAKE IN EPIDEMIOLOGIC ANALYSES

When relationships with disease are analyzed, nutritional factors may be examined in terms of absolute amount (crude intake) or in relation to total caloric intake. The analytic approach depends on both the nature of the biologic relationship and the public health considerations. If a nutrient is metabolized in approximate proportion to total caloric intake (such as the macronutrients and some vitamins), nutrient intake is most likely biologically important in relation

to caloric intake. To the extent that energy intake reflects body size, adjustment for total energy intake is usually appropriate as an absolute amount of a specific nutrient tends to have less of an effect for a larger, higher energy-consuming person than for a smaller person. In some situations, we may be unsure whether it is the absolute amount of a nutrient or the amount in relation to total caloric intake that is most biologically relevant. (Of course, the biologically relevant relationship with caloric intake may actually be complex and nonlinear.) If a nutrient selectively affects an organ system that is uncorrelated with body size (e.g., the central nervous system), or if physical activity does not affect its metabolism, absolute intake may be most relevant.

It is interesting to note that if absolute nutrient intake rather than intake in relation to calories is biologically most relevant, caloric intake should be associated with disease, as intakes of virtually all nutrients are positively correlated with total caloric intake. For example, if higher absolute intake of a nutrient is a determinant of disease, then those who consume more total food due to being large, active, or metabolically inefficient should be at higher risk of disease. Conversely, lack of association between total energy intake and disease can be taken as evidence against the importance of absolute nutrient intake, but not against the importance of nutrient composition of the diet.

Because a person's long-term total caloric intake is largely determined by body size, physical activity, and metabolic efficiency, even relatively small changes in caloric intake cannot be made unless changes in weight or physical activity also occur. In the absence of such changes, therefore, most alterations in absolute nutrient intake must be accomplished by changing the composition of the diet rather than the total amount of food. For this reason, dietary recommendations should be made in reference to total caloric intake (Hegsted, 1985); for example, it has been suggested that we reduce total fat intake from 40 to 30 percent of total caloric intake. Therefore, from a practical or public health standpoint, nutrient intake in relation to total caloric intake (i.e., the qualitative aspect of diet) is most relevant.

Because reports of diet and disease relationships are often made without adjustment for total caloric intake, the implications of this approach are discussed first. Because body size, physical activity, and metabolic efficiency contribute to the variation of specific nutrient intakes, associations between nutrients and diseases that are actually independent of these factors are weakened. That is, differences in the levels of these factors cause variation in energy intake and, secondarily, variation in intake of specific nutrients that may be extraneous or irrelevant to occurrence of disease. For example, tall and physically active women tend to have high absolute fat intakes on the basis of their size and exercise level alone. If the relevant exposure for breast cancer risk is fat intake independent of body size and physical activity (i.e., the fat composition of the diet), failure to account for these factors results in misclassification of exposure that is likely to be largely random. To partially address this issue, nutrient intakes are often divided by a measure of body size, such as intake per kilogram (Sopko et al., 1984). Unfortunately, it is seldom possible to account for the effects of physical activity and metabolic efficiency in a similar manner because these are difficult to measure.

Table 11–2. Correlations (Spearman r) between intakes of specific nutrients

	Protein	Total fat	Saturated fat	Polyunsaturated fat	Total carbohydrate	Crude fiber	Vitamin A	Sucrose	Vitamin B_6	Vitamin C	Cholesterol	Alcohol
Protein	—	—	—	—	—	—	—	—	—	—	—	—
Total fat	0.44											
	−0.13	—	—	—	—	—	—	—	—	—	—	—
	−0.14											
Saturated fat	0.42	0.93										
	−0.10	0.74	—	—	—	—	—	—	—	—	—	—
	−0.10	0.74										
Polyunsaturated fat	0.28	0.73	0.52									
	−0.08	0.56	0.02	—	—	—	—	—	—	—	—	—
	−0.11	0.56	0.02									
Total carbohydrate	0.39	0.56	0.47	0.41								
	−0.34	−0.48	−0.43	−0.15	—	—	—	—	—	—	—	—
	−0.23	−0.52	−0.57	−0.14								
Crude fiber	0.36	0.01	−0.06	0.12	0.45							
	0.32	−0.44	−0.51	−0.07	0.35	—	—	—	—	—	—	—
	0.24	−0.45	−0.51	−0.08	0.45							
Vitamin A	0.45	0.08	0.03	0.10	0.66	0.66						
	0.40	−0.33	−0.38	−0.08	0.13	0.64	—	—	—	—	—	—
	0.37	−0.32	−0.36	−0.08	0.19	0.61						

Sucrose	0.21	0.53	0.46	0.37	0.18	0.18	0.14					
	−0.54	−0.18	−0.19	−0.13	0.60	−0.17	−0.20	—	—	—	—	—
	−0.45	−0.21	−0.25	−0.11	0.56	−0.05	−0.12					
Vitamin B_6	0.56	0.15	0.14	0.12	0.65	0.65	0.57	0.22				
	0.44	−0.49	−0.39	−0.22	0.26	0.61	0.53	−0.23	—	—	—	—
	0.41	−0.48	−0.38	−0.22	0.33	0.59	0.59	−0.16				
Vitamin C	0.34	0.002	−0.05	−0.00	0.61	0.61	0.13	0.13	0.55			
	0.32	−0.47	−0.46	−0.14	0.17	0.62	0.60	−0.12	0.53	—	—	—
	0.26	−0.45	−0.44	−0.15	0.34	0.59	0.58	−0.04	0.51			
Cholesterol	0.60	0.51	0.55	0.26	0.05	0.05	0.25	0.18	0.24	0.12		
	0.58	0.16	0.22	−0.02	−0.36	−0.00	−0.17	−0.34	0.09	0.11	—	—
	0.50	0.19	0.26	−0.04	−0.32	−0.07	0.13	−0.30	0.05	0.03		
Alcohol	−0.03	0.04	0.12	−0.01	−0.15	−0.14	−0.03	−0.06	−0.06	−0.06	−0.04	
	−0.16	−0.12	0.06	−0.13	−0.50	−0.19	−0.10	−0.14	−0.16	−0.14	−0.15	—
	−0.13	−0.12	0.05	−0.13	−0.54	−0.18	−0.07	−0.21	−0.14	−0.12	−0.12	
Energy	0.59	0.86	0.81	0.59	0.82	0.24	0.25	0.71	0.40	0.19	0.47	0.15

For each comparison, top r value is for crude nutrient intake, middle value is for nutrient density, and the bottom value is for calorie-adjusted (using regression analysis) nutrient intake. Data are based on the individual means of 28 days of diet recording by each of 194 women.

Table 11–3. Correlations (Pearson r) between total caloric intake, crude nutrient intake, nutrient densities, and calorie-adjusted nutrient intakes

Nutrient	Calories vs. crude nutrient	Nutrient density[a] vs crude nutrient	Calorie-adjusted vs crude nutrient[b]	Calories vs nutrient density[a]	Calories vs calorie-adjusted[b]	Nutrient density[a] vs calorie-adjusted[b]
Protein	0.60	0.31	0.80	−0.57	0.000	0.82
Total fat	0.88	0.52	0.48	0.05	0.000	0.99
Saturated fat	0.81	0.66	0.58	0.09	0.000	0.99
Polyunsaturated fat	0.64	0.77	0.77	−0.01	0.000	0.99
Sucrose	0.72	0.91	0.70	0.33	0.000	0.93
Cholesterol	0.47	0.70	0.88	−0.31	0.000	0.95
Carbohydrate	0.86	0.72	0.52	0.26	0.000	0.97
Fiber	0.34	0.82	0.94	−0.26	0.000	0.97
Vitamin A-ns	0.34	0.84	0.94	−0.23	0.000	0.97
Vitamin A-ws	0.25	0.90	0.97	−0.19	0.000	0.98
Vitamin B_6-ns	0.46	0.76	0.89	0.22	0.000	0.97
Vitamin B_6-ws	0.15	0.98	0.99	−0.06	0.000	0.99
Vitamin C-ns	0.28	0.88	0.96	−0.20	0.000	0.98
Vitamin C-ws	0.15	0.96	0.99	−0.13	0.000	0.99

Data are based on the individual means of 28 days of dietary recording by each of 194 women and on four 1-week diet records. All values were transformed using natural logarithm to improve normality.

[a]Nutrient density is the nutrient divided by calories.

[b]Calorie-adjusted using regression analysis.

ns = without supplements, ws = with supplements.

When total caloric intake is associated with disease, the interpretation of individual nutrient intake is complex, and the consequences of failing to account for energy intake may be far more serious. As has been pointed out by Lyons and colleagues (1983), specific nutrients tend to be associated with disease simply on the basis of their correlation with caloric intake. For example, in nearly every study of diet and coronary heart disease, subjects who subsequently develop disease have lower total caloric intake on the average than those who remain free of disease (Morris et al., 1977; Garcia-Palmieri et al., 1980; Gordon et al., 1981; McGee et al., 1984; Kromhout and Doulander, 1984). As a result, intakes of many individual specific nutrients also tend to be lower among cases than among noncases.

These relationships are illustrated in analyses of prospectively collected data on dietary intake and coronary heart disease incidence among men living in Honolulu, Puerto Rico, and Framingham (Gordon et al., 1981). Among the Honolulu men, for example, the crude intakes of nine of 11 nutrients (including total calories) were lower among those with subsequent coronary heart disease, and for two nutrients there was no difference (Table 11–4). In this situation, it is helpful to consider possible reasons for an observed difference in caloric intake between men who developed coronary disease and those who remained free of disease. Difference in body size is an unlikely explanation as men who subsequently develop coronary heart disease tend, if anything, to weigh more than those who do not. Variation in level of metabolic efficiency is usually impossible to eliminate as an explanation. On the other hand, several investigators have

Table 11–4. Age-adjusted means of crude nutrient intakes and nutrient intakes as a percentage of total calories according to subsequent coronary heart disease (CHD) death or myocardial infarction (MI)

	Crude intakes		Intakes as % of calories	
	No CHD (n = 7,008)	MI or CHD death (n = 164)	No CHD (n = 7,008)	MI or CHD death (n = 164)
Total calories (kcal)	2319	2149[b]		
Total protein (g)	95	93	16.6	17.4[a]
Total fat (g)	87	86	33.4	35.6[a]
Saturated fat (g)	32	31	12.3	12.9[a]
Monounsaturated fat (g)	33	32	12.8	13.6[b]
Polyunsaturated fat (g)	16	16	6.0	6.7
Total carbohydrate (g)	264	242[b]	46.2	45.4
Sugar (g)	46	46	7.9	8.2
Starch (g)	165	151[b]	29.2	28.5
Other carbohydrate (g)	52	45[a]	9.1	8.7
Cholesterol (mg)	555	530		
Alcohol (g)	14	5[b]	3.8	1.7[b]

[a] $p < 0.05$.

[b] $p < 0.01$.

Data are based on a cohort of 7172 Honolulu men aged 46–64 years initially free of coronary heart disease. From tables 4 and 8 of Gordon et al., 1981.

clearly demonstrated that decreased physical activity is associated with an increased risk of coronary heart disease (Morris et al., 1977; Paffenbarger et al., 1978). Although differences in physical activity provide the most likely explanation for the low caloric intake associated with coronary heart disease, this explanation has not been universally appreciated (Garcia-Palmieri et al., 1980; McGee et al., 1984). Thus, an appropriate interpretation of the inverse association between total caloric intake and risk of coronary heart disease is not that one should increase food intake to avoid a myocardial infarction, but rather that an increase in physical activity may reduce the risk of disease. This example, incidentally, illustrates the need to be guided by an understanding of biologic relationships when interpreting statistical associations to avoid absurd conclusions.

Because variation in caloric intake between persons largely reflects physical activity, size, and metabolic efficiency, an association between a specific nutrient and disease is not likely to be of primary etiologic importance if that association is simply the result of a difference in caloric intake. For this reason, Morris and co-workers (1977) have pointed out that it "would not be instructive to present data relating crude nutrient intakes with disease in a situation in which caloric intake has an important relationship with the outcome."

It is, of course, possible that overeating or undereating (caloric excess or deficiency) is a primary cause of a disease. In this situation, nutrients that contribute to calories (proteins, fats, carbohydrates, and alcohol) might be considered as the primary exposures that lead to increased caloric intake, which in turn causes

disease. It could be argued that adjustment for caloric intake in this situation would represent "over-control" of a variable in the causal pathway. Before attributing an effect to a specific nutrient, however, the burden is on the epidemiologist to demonstrate that the association of this nutrient with disease is independent of caloric intake. For example, perhaps excessive caloric intake increases the risk of colon cancer and dietary fat is associated with this disease because of its high caloric content. Before implicating fat per se as a specific cause, however, it would be essential to demonstrate that this effect is not shared by protein or carbohydrate when these are eaten in equicaloric amounts. Otherwise, a reduction in the fat content of the diet that was substituted by an increase in carbohydrate or protein on a calorie-by-calorie basis would have no effect on disease occurrence: this would only happen when the total caloric intake was also changed. The desirability of relating nutrient intake to total caloric intake has been discussed in a thoughtful correspondence with respect to studies of coronary heart disease (Shekelle et al., 1985; Kushi et al., 1985). Recognizing the need to adjust for the effect of total food consumption, a number of investigators have employed "nutrient densities" to control for the effect of total caloric intake.

Analyses of Diet–Disease Relationships by the Use of Nutrient Densities

Nutrient densities are computed by dividing nutrient values by total caloric intake; they provide a convenient way to describe foods or diets. An analogous approach for macronutrients is to express intake as a percentage of total caloric intake; for purposes of discussion, both approaches are referred to as nutrient density. Nutrient density has the appeal of simplicity and practicality; unfortunately, this is actually a complex variable with a generally obscure meaning when used to address diet–disease relationships. Such a variable has two components: the nutrient intake and the inverse of total caloric intake. The relative contributions of nutrient intake and total caloric intake to between-person differences in nutrient density are related to the ratio of their variability. Thus, as the between-person variation in the specific nutrient intake decreases, the nutrient density value approaches the inverse of caloric intake (multiplied by a constant).

When energy intake is unrelated to disease, dividing nutrient intakes by total calories may contribute to the desired effect of reducing variation in nutrient intake that is due to differences in size, activity, and metabolic efficiency. The division, however, also can create unwanted variation. Particularly when a specific nutrient has a weak correlation with total energy intake or has a low variability, dividing by total calories creates a variable that is, in fact, highly related to the factor whose effect we wish to remove, that is, caloric intake. In addition, methodologic error in measuring total energy intake could potentially contribute to variation in nutrient density as a result of this division. The basic principle involved is that dividing by a variable does not necessarily remove or "control for" the effect of that variable.

As with crude nutrient intakes, the use of nutrient densities has more serious

implications when total energy intake is itself associated with disease. Because a nutrient density variable contains the inverse of energy intake as a component, nutrient densities tend to be associated with disease in the direction opposite to that of total caloric intake, even when the nutrient itself has no association with disease independent of energy intake. The data of Gordon and colleagues (1981), which present nutrient intakes as percentages of total calories, again illustrate this point. Because coronary heart disease is associated with low caloric intake, nutrient densities (or intakes as percentages of total calories) tend to be positively associated with disease (Table 11–4). In this instance dividing by total calories has changed the direction of association with coronary heart disease for protein and total, saturated, monounsaturated, and polyunsaturated fat, and four of these differences become statistically significant. Differences between cases and controls have essentially disappeared for the three measures of carbohydrate intake that were statistically significant in the crude analysis.

In some studies, the reason for a difference in energy intake is obscure. In a carefully conducted Canadian case-control investigation of large bowel cancer by Jain and co-workers (1980), cancer patients reported higher caloric intake than did controls but did not weigh more than the controls (see crude intakes, Table 11–5). In addition, cancer cases consistently reported higher intakes of fat

Table 11–5. Case minus control differences in crude and nutrient density intakes expressed as a percentage of case value

	Case-control difference (%)							
	Crude intake (original analysis)				Nutrient density intake (recalculation)			
	Males		Females		Males		Females	
	Colon	Rectum	Colon	Rectum	Colon	Rectum	Colon	Rectum
Calories								
Neighborhood controls	9[b]	7	12[c]	17[c]				
Hospital controls	1	9[a]	6[a]	11[b]				
Total fat								
Neighborhood controls	8[b]	6	15[c]	22[c]	0	−1	4	7
Hospital controls	2	11[a]	10[a]	15[b]	2	2	4	5
Saturated fat								
Neighborhood controls	13[b]	8	16[c]	27[c]	4	1	5	12
Hospital controls	6	13[a]	9[a]	19[c]	5	4	2	8
Crude fiber								
Neighborhood controls	−5	−3	1	2	−15	−11	−12	−17
Hospital controls	−2	5	7	5	−3	−5	1	−7
Vitamin C								
Neighborhood controls	−3	4	−4	0	−13	−3	−18	−20
Hospital controls	−2	6	−2	−3	−3	−6	−9	−16

[a] $p < 0.05$.

[b] $p < 0.01$.

[c] $p < 0.002$.

Data are calculated from table 5 of a case-control study of colon and rectal cancer conducted among Canadian men and women between 1976 and 1978 (Jain et al., 1980). No tests of statistical significance available for nutrient density data.

From Willett and Stampfer, 1986.

than did noncases, which was interpreted to "support the hypothesis that high dietary fat intake is causally associated with cancer of the colon and rectum." In interpreting these findings, it is again useful to consider possible explanations for the difference in caloric intake between cases and controls. It seems unlikely that higher physical activity by cancer cases would explain their higher energy intake; indeed, data to the contrary have been published (Garabrant, 1984; Vena et al., 1985). We cannot dismiss the possibility that cases have a metabolic abnormality that renders them less efficient in their utilization of food energy. For example, it is conceivable that subjects who absorb food poorly present more substrate to their fecal flora, which metabolize this to carcinogenic substances, and thus increase the risk of large bowel cancer. The case-control difference in caloric intake is unlikely to be the result of simple overeating as cases were not more obese than controls, even several years before diagnosis. (The lack of association with relative weight found by Jain and co-workers has also been observed in prospective incidence data; see Sidney et al., 1986.)

In addition to biologic factors, recall bias cannot be dismissed as an explanation for the findings of Jain and colleagues (1980) as this was a case-control study. Whatever biologic or methodologic factors contribute to the association of caloric intake with colon cancer, any differences in the intake of specific nutrients between cases and controls that result from the strong association of caloric intake with cancer must be regarded as secondary. For illustrative purposes, we recalculated intakes as nutrient densities instead of crude values as originally presented; the findings were dramatically altered (Table 11–5). The association with total fat intake essentially disappears for men and is largely eliminated for women. On the other hand, strong inverse associations are seen for fiber and vitamin C intakes expressed as nutrient densities, which had no association with cancer in the crude analysis. This nutrient density analysis, however, overstates the protective association of fiber and vitamin C and underestimates the effect of fat, because dividing by caloric intake produces inverse associations even when these nutrients are not independently associated with disease. Jain and colleagues recognized the potential for confounding by total caloric intake and stated that the effects of fat and total caloric intake were difficult to separate as it was not possible to enter both simultaneously in a logistic model because of their high correlation. Instead they considered fat rather than calories to be the primary factor due to its stronger (albeit slightly) association with cancer and the findings of previous animal studies. On the basis of the data presented in the original article, the findings for fat are difficult to interpret as the positive association of this and other nutrients is overstated in crude analyses and understated in nutrient density analyses. [This analysis is presented as an example; these authors have published additional data on this topic (Howe et al., 1986) and provided further personal communications demonstrating that, in their study, saturated fat is positively associated with risk of colon cancer independent of energy intake, but fiber intake is not independently related to risk.] If it could be demonstrated that the positive association between caloric intake and colon cancer incidence is related to a real difference in metabolic efficiency between cases and noncases rather than to methodologic bias, this would be an important increment in knowledge, even though the association

would not represent a primary etiologic effect of diet and, therefore, have no implications for nutritional advice.

Alternative Analytic Approaches

For reasons discussed previously, it is usually desirable in epidemiologic analyses to employ a measure of nutrient intake that is independent of total caloric intake, particularly when caloric intake is associated with disease. In this section four analytic strategies and their relationships are considered: the "energy-adjusted" method, the standard multivariate method, the "energy decomposition" method, and the multivariate nutrient density method.

Energy-adjusted Method

"Energy-adjusted" nutrient intakes are computed as the residuals from the regression model with total caloric intake as the independent variable and absolute nutrient intake as the dependent variable. Because residuals have a mean of zero and include negative values, they do not provide an intuitive sense of actual nutrient intake. It may, therefore, be desirable to add a constant; logical choices are the expected nutrient intake for the mean caloric intake of the study population at hand or a round number for energy intake near the population mean (Fig. 11–5). If the usual assumptions for regression analysis are met, these calorie-adjusted nutrient intakes are not associated with caloric intake.

When total energy intake is an important predictor of disease, total caloric intake should be included in the model with the nutrient calorie-adjusted term (model 2, Table 11–6). This approach has theoretical advantages over entering only calorie-adjusted nutrient into a model, as the random error (and width of confidence limits for the effect of the nutrient) is reduced if caloric intake has an important association with disease independent of nutrient intake. Another advantage of using model 2 rather than model 1 is that the full effect of total caloric intake is observed. To illustrate adjustment for total energy intake by regression analysis, we have used daily intakes of total calories and total fat based on the means of four 1-week diet records kept by each of 194 women, as described previously (Willett et al., 1985). With 28 days of recording per subject, the effect of day-to-day variation has been sufficiently dampened such that these values can be assumed to reasonably represent each subject's long-term intake. The unadjusted intake of total fat has a reasonably wide distribution (mean = 68.9, 1 standard deviation = 17.0 g per day; Fig. 11–6). Because total fat and total caloric intake are highly correlated ($r = 0.86$), adjustment for total caloric intake reduces the variation in fat intake substantially (mean = 68.9, 1 standard deviation = 8.7 g per day; shaded area in Fig. 11–6). Nevertheless, the degree of variation remaining is realistic in relation to current dietary recommendations; the tenth percentile (median of the lowest quintile) represents 33 percent of calories from fat, and the ninetieth percentile represents 44 percent of calories from fat. Thus, sufficient variation exists in this population to test the effectiveness of a 25 percent reduction in the proportion of calories accounted for by fat, as recommended by the National Heart, Lung, and Blood Institute (NIH Consensus Conference, 1985).

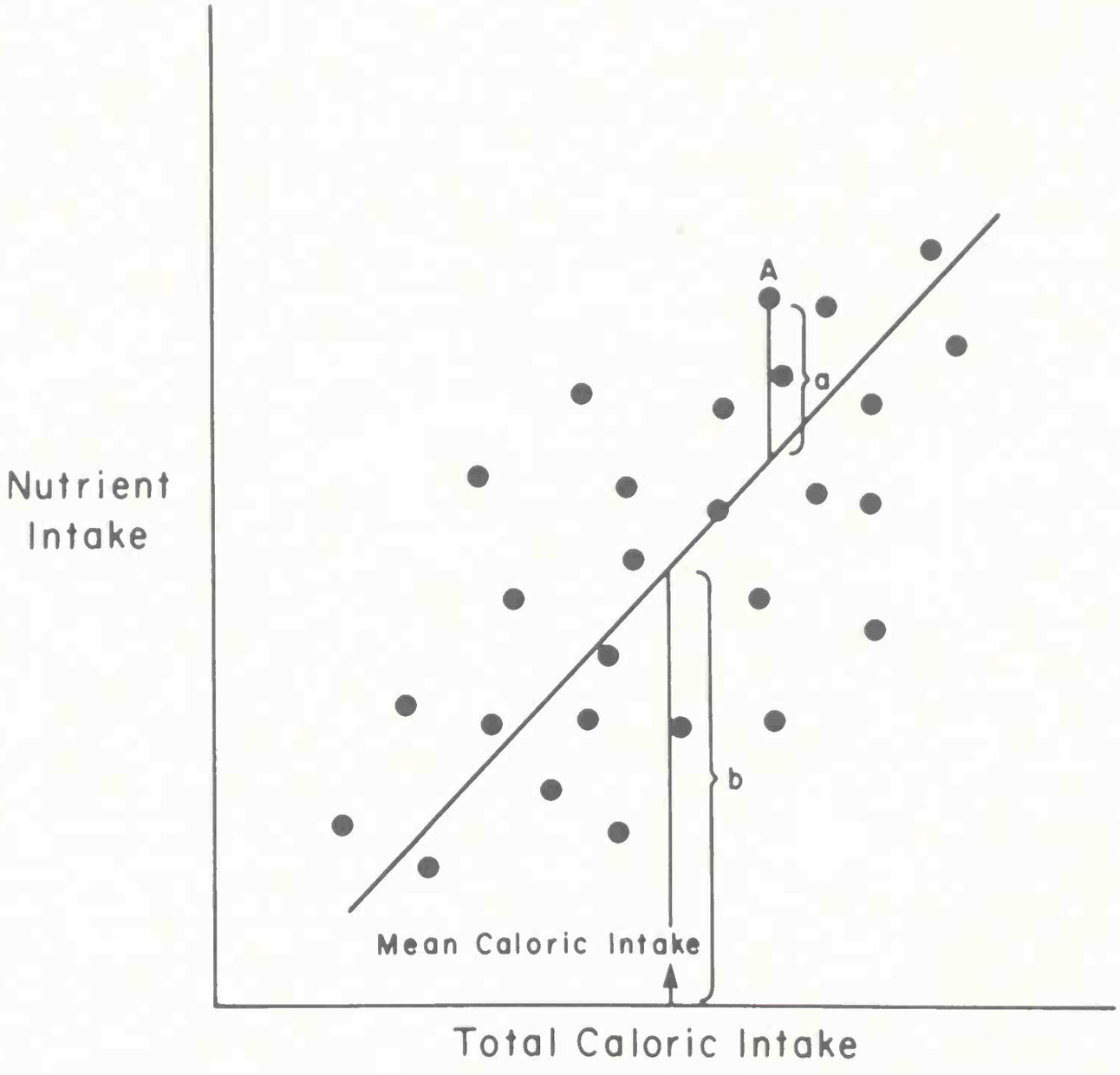


Figure 11–5. Calorie-adjusted intake = a + b, where a = residual for subject from regression model with nutrient intake as the dependent variable and total caloric intake as the independent variable and b = the expected nutrient intake for a person with mean caloric intake. (From Willett and Stampfer, 1986; reproduced with permission.)

Shekelle and Nichaman (1987) have noted the desirability of transforming nutrient variables (which are usually skewed with a long tail to the right), to improve normality before adjustment using regression analysis. It has also been pointed out (M. Maclure, personal communication) that the use of highly skewed nutrient variables usually results in those with higher caloric intakes having large residuals and thus more likely to be in the lowest and highest categories of calorie-adjusted intakes. As with any analysis, it is important to make appropriate transformations ($\log_e$ is frequently appropriate) so that model assumptions are not seriously violated.

Shekelle and Nichaman (1987) have also calculated calorie-adjusted nutrient intakes using regression analysis and examined their correlations with nutrient densities; these correlations were consistently high (greater than 0.90). On this basis they suggested that, although the calorie-adjusted values were theoretically preferable, the use of nutrient densities may not necessarily lead to materially

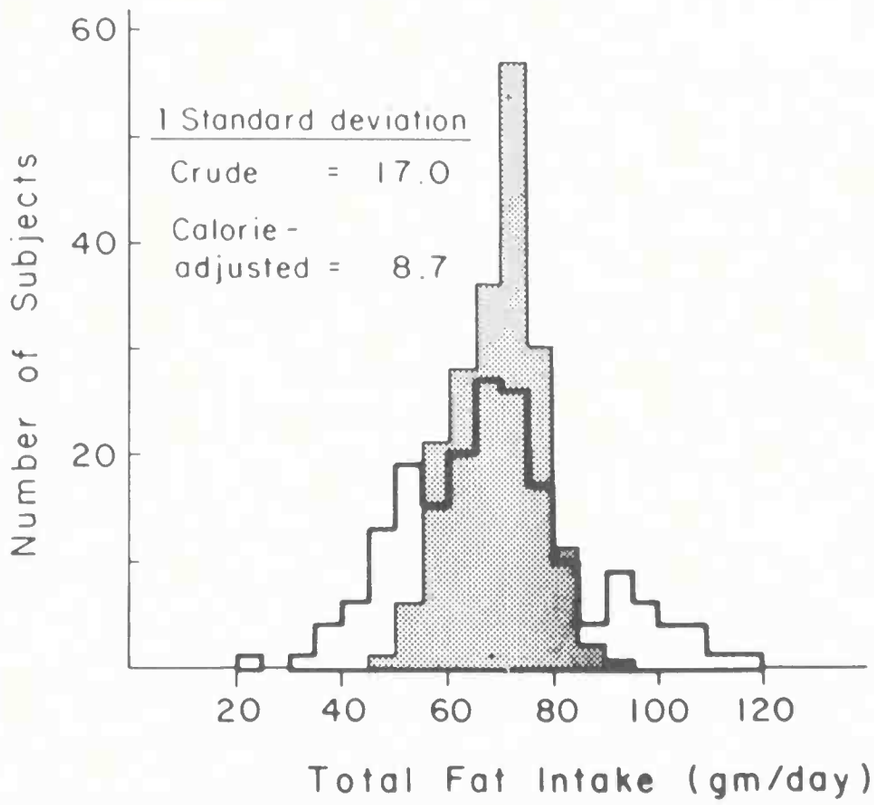


Figure 11–6. Distribution of total fat intake with *(shaded area)* and without *(dark line)* adjustment for total caloric intake. Data are based on four 1-week diet records completed by 194 Boston-area women aged 34 to 59 years . Calorie-adjusted values were calculated as described in the text, with residuals computed in the $\log_e$ scale to improve normality. (From Willett and Stampfer, 1986, reproduced with permission.)

different conclusions in epidemiologic analyses. We observed similarly high correlations in our own data (see Table 11–3), but were not confident that the difference between calorie-adjusted values and nutrient densities might not be important in some instances, as the major concern is the potential for confounding by calories. As shown in Table 11–3, the correlations between calorie-adjusted values and calories is, by definition, zero. Several nutrient density values, however, particularly for those nutrients less strongly correlated with total energy intake, were moderately correlated with caloric intake. For example, the correlation of calories with protein/calories was −0.57 and with fiber/calories was −0.26, meaning that the potential for confounding by total calories is substantial if energy intake is related to disease. Because the magnitudes of these correlations are partly related to behavior rather than by laws of nature, these correlations vary from one group to another. The degree of confounding depends on the strength of association between energy intake and disease as well as nutrient intake and disease; therefore, it seems that distinction between nutrient densities and regression-adjusted values is likely to have practical importance in some, although certainly not all, instances. It appears that this distinction is of major importance for fiber intake in the case-control study by Jain and co-workers (1980) because the nutrient density data in Table 11–5 imply a protective association whereas the regression-adjusted analysis reportedly did not. An examination of relative risks computed using both nutrient densities and calorie-adjusted values would be worthwhile in other data sets in which energy intake is associated with disease.

The Standard Multivariate Method

The residual approach of calorie adjustment (model 1, Table 11–6) is similar, although not identical, to including both caloric intake and absolute nutrient intake as terms in a multiple regression model with disease outcome as the dependent variable (model 3). The coefficient for the nutrient term in this multivariate model (b_4) is identical to that for the calorie-adjusted nutrient term in a univariate model (b_1). With total caloric intake also in the calorie-adjusted model (model 2), the standard error as well as the coefficient for the calorie-adjusted nutrient term (b_1) is identical to that for the nutrient term (b_4) in the standard multivariate model with nutrient and calories.

Although similar in some respects, the standard multivariate model (model 3) creates complexities in interpretation not shared by model 2. If calories and nutrient are simply entered as separate terms, the coefficient for calories (b_3) represents calories independent of the specific nutrient, which may have a meaning distinctly different than total energy intake. For example, if fat is the specific nutrient in model 3, then the term for calories has the meaning of carbohydrate plus protein (not considering the possible contribution of alcohol). Thus the inclusion of a specific nutrient together with calories in a model fundamentally changes the biologic meaning of calories. The coefficient for calories in this model (b_3) may, therefore, fail to attain significance when total energy intake, in fact, has a significant and important relation with disease. In contrast, the two terms in model 2 clearly address two distinct and clear questions: is total energy intake associated with disease? and, is the nutrient composition of the diet related to disease? The simplicity and clear meaning of the calorie-adjusted intakes also make them attractive for bivariate analyses and data presentation.

Some authors have voiced concern over the simultaneous inclusion of strongly correlated variables in the same model, which will frequently occur using the standard multivariate model in nutritional studies. McGee and colleagues (1984) have noted that widely divergent results are obtained when highly correlated variables are entered in multiple logistic models using various inclusion criteria, and they suggest that variables with correlation coefficients of more than approximately 0.60 not be simultaneously included. To the extent that high colinearity is a concern, the use of calorie-adjusted nutrient intakes reduces this

Table 11–6. Alternative models for addressing the correlations of specific nutrient intakes with total energy intake in epidemiologic analyses

Model 1	Disease = b_1Nutrient residual[a]
Model 2	Disease = b_1Nutrient residual + b_2Calories
Model 3	Disease = b_3Calories + b_4Nutrient
Model 4	Disease = $b_5Cal_{Nutrient}$[b] + b_6Cal_{other}[c]
Model 5	Disease = b_7Nutrient/Calories + b_8Calories

[a] "Nutrient residual" is the residual from the regression of a specific nutrient on calories.

[b] $Cal_{nutrient}$ represents calories provided by the specific nutrient.

[c] Cal_{other} represents calories from sources other than the specific nutrient.

problem. The issue of colinearity is, however, better viewed in biologic rather than purely statistical terms. The first problem created by including two strongly correlated variables in a model is that their meaning may change in a manner that is not readily appreciated. The example of calories and fat is noted above and the problem of height becoming a measure of body composition when weight is included in a model was noted in Chapter 10. The second problem resulting from colinearity is that one or both variables may have a markedly reduced degree of variation when they are entered simultaneously. Rather than using any arbitrary level of correlation as a criteria for unacceptable colinearity, however, a more informative approach is to examine the degree of residual variation in the variables of interest and judge whether the remaining differences among individuals are worthy of study. For example, even though the correlation between total energy and total fat intakes was 0.86 in the data displayed in Figure 11–5, the residual variation in fat after adjusting for total calories was found to be of potential interest. If the residual variation in the variable of interest, such as fat intake, is not large enough to be informative, the issue is not statistical but relates to the nature of the study population; the only solution is to find another population with a wider residual variation. When faced with highly correlated nutritional variables, models 1 and 2 allow a clear specification of the meaning of the variables as well as the opportunity to evaluate the residual variation in nutrient intake directly.

The "Energy Decomposition" Method

Howe and colleagues (1986) have presented an alternative multiple logistic regression model in which they entered separate terms for energy from a specific macronutrient, such as fat, and for energy not-from-fat, meaning protein, carbohydrate and alcohol (Table 11–6, model 4). In this model, the coefficient for the specific macronutrient (b_5) represents full effect of the nutrient unconfounded by other sources of energy (b_6), but this model does not directly address the question of whether energy from the specific nutrient has an association with disease not shared by other sources of energy, in other words, independent of total energy intake. To address this issue would require determining whether the magnitude of the coefficient for the specific nutrient (b_5) was actually different from the coefficient for other sources of energy (b_6), in other words, the appropriate focus should be b_5–b_6. It is not adequate merely to note that the nutrient coefficient (b_5) is significant, whereas b_6 is not; even when all sources of energy have the same relation with disease on a calorie-for-calorie basis, the coefficient for other calories (b_6) might not be significant simply because of low between-person variation in this factor. In fact, it can be shown that the difference in these coefficients (b_5–b_6) and the standard error of this difference are identical to the coefficient and standard error for the nutrient residual (b_1) in model 2.

Although the "energy decomposition" model may provide insight in some instances, its coefficients may be misleading unless interpreted with care, particularly when total energy intake has a noncausal relationship with disease. For example, use of this model in the example of coronary heart disease noted previously could easily indicate a protective association for fat intake as its effect

would still be confounded by total energy consumption secondary to differences in physical activity. An additional limitation of this model is that it cannot be readily extended to nutrients that do not contribute to energy intake.

Multivariate Nutrient Density Model

Another approach, used in the article by Jain and colleagues (1980) to examine the effect of fiber, is to compute the nutrient density, then enter both this and total energy in a multiple logistic regression model (model 5). Although this method is somewhat indirect, it does control for confounding by energy intake; the coefficient for the nutrient density term (b_7) represents the relation of the nutrient composition of the diet with disease, holding total energy intake constant. This model overcomes the statistical problem associated with the use of the nutrient density alone, while retaining its attractive features of general recognition and intuitive interpretation as a measure of dietary composition. The coefficient for calories in this model (b_8) will generally be interpretable as representing the effect of calories in the usual biological sense because nutrient densities are not inherently part of or highly correlated with total energy intake.

The multivariate nutrient density model may be particularly advantageous when body size (and thus total energy intake) varies greatly among subjects because models 2 and 3 imply that the nutrient residual has a similar effect for subjects with high and low energy intake. That a given increment in nutrient intake would have the same effect in a very small subject (with low energy intake) as in a very large subject (with high energy intake) is not plausible. (In principle, this complexity should be addressed in models 2 and 3 by using an interaction term; however, the use of log-transformations, as discussed earlier, will tend to minimize this problem.) Among specific age-sex groups of human populations, variation in size and total energy is not great, so that this is usually not a major issue. However, among other species, such as dogs (Sonnenschein et al., 1987), body size can vary more than 10-fold, making the use of the multivariate nutrient density model particularly attractive.

The Energy Determinant Method An alternative approach, in theory, would be to include the major determinants of energy intake (body size, physical activity, and metabolic efficiency) as separate variables in a multivariate model. Unfortunately, measurements of these variables are usually not available in epidemiologic studies. It could be informative, however, to include as many of these variables as possible along with total energy intake as independent variables. Because energy intake and disease outcome may differ in their relationships with body components such as lean mass and fat, it would be desirable to include both height and a measure of fatness uncorrelated with height as separate terms in a multivariate model. The residual of weight on height, computed as described previously for calorie-adjusted nutrients, provides a measure of weight uncorrelated with height. Modeled in this way, height represents lean body mass as it has a linear relationship with total body water in adults (Mellits and Cheek, 1970), and weight independent of height primarily represents fat in

middle-aged and older subjects. The interpretation of energy intake as an independent variable depends on which other terms are included. If one assumes a steady state of energy balance, energy intake adjusted for body size and physical activity has the meaning of metabolic efficiency. When adjusted for height and weight only, total energy intake has the meaning of physical activity and metabolic efficiency combined. In real applications, these interpretations must be tempered with the knowlege that physical activity is probably measured only crudely and, in case-control studies, that energy intake may also include an overall bias in the measurement of dietary intake.

More Complex Models. In the analytic approaches discussed previously, only one macronutrient at a time was considered in addition to total energy intake. In principle, these approaches could be extended to include other nutrients as well. For example, using the energy-adjustment approach (model 2), one could compute calorie-adjusted residuals for both protein and fat and include both along with total calories in the same model, or one could use the energy decomposition method to enter energy from fat, protein, and carbohydrate as three separate terms. Whether sufficient independent variation in macronutrient intake exists in typical data sets to produce stable coefficients using these more complex models is not clear at this time. Clearly, the inclusion of additional nutrient terms to these models should be done with caution as the interpretation of even the two-variable models can be complex.

In the energy-adjustment method (models 1 and 2), further question about the constituents of macronutrients could be examined by nesting one residual within another. For example, if we are interested in the effect of saturated fat intake independently of total fat intake, a residual could be computed with saturated fat intake as the dependent variable and calorie-adjusted total fat intake as the independent variable. The model would then include calories, calorie-adjusted total fat, and total fat-adjusted saturated fat. The latter term would indicate whether saturated fat was uniquely associated with disease apart from its contribution to overall fat intake. In the energy decomposition model, separate terms would be entered for calories from saturated fat, calories from other types of fat, and calories from sources other than fat. To evaluate the specific effect of saturated fat in this model, the difference between the coefficients for the saturated fat and other types of fat terms would be examined.

As in most epidemiologic analyses, the possibility of interaction should be considered. In the energy-adjustment model (model 2), an interaction could be represented by a cross-product term for total calorie intake and the calorie-adjusted nutrient intake. Such an interaction is plausible as a specific difference between observed and predicted nutrient intake is likely to have less of an effect for a person with large energy intake than for a person with small energy intake. In practice, such an interaction is made less likely by the usual use of the log scale to avoid heteroscedasticity. Although the presence of a strong interaction would have the undesirable effect of complicating the model, this possiblity should be entertained, by evaluating the interactive term, rather than ignored.

Implications for Food-Frequency Questionnaire Data

The preceding discussion assumes that accurate, quantitative data are available for analysis. Because of the need for rapid, inexpensive methods to assess long-term intake in large numbers of subjects many epidemiologists use simple (Todd et al., 1983) or semiquantitative food-frequency questionnaires that are clearly less than perfectly accurate. Typically, these questionnaires are used to compute estimates of nutrient intake with the primary objective of ranking or categorizing subjects rather than providing precise quantitative measures of absolute intake.

The meaning of energy intake computed from such questionnaires may be less clear than that from more quantitative methods. To the extent that subjects with higher caloric intakes simply consume larger portion sizes rather than more food items, nutrient intakes may be inherently adjusted for total caloric intake. This adjustment, however, is likely to be only partial at most as many food items (e.g., eggs, bread, and apples) come in predetermined units. With any method of assessing energy intake, subjects may either overestimate or underestimate their overall intake. Although direct evidence is not available, it is possible that these tendencies are greater with food-frequency questionnaires.

Although energy intake data from food-frequency questionnaires may be imperfect and thus not fully represent the effects of body size, activity, metabolism, and energy balance, it would still be appropriate to use this measure for the computation of energy-adjusted intakes as described previously. To the extent that this adjustment also reduces between-person variation due to general overreporting or underreporting of intake, a further gain in accuracy may be obtained in some instances (see Chapter 8).

Although adjustment for total caloric intake based on a questionnaire should reduce confounding by caloric intake, it must be recognized that both the nutrient and caloric intake are imperfectly measured so that control of confounding may not be complete (Greenland, 1980). Data from a validation study can be extremely useful to evaluate the degree to which confounding has been controlled. For example, within the Nurses' Health Study a reasonably strong association was observed between risk of a certain disease and intake of both calories and total fat. The association with calorie-adjusted fat was slightly less strong, but still potentially important. We were concerned, however, that the effect of calorie-adjusted fat might be due to residual confounding by total energy intake, which was measured imperfectly by the questionnaire. Therefore, we examined the correlation between calorie-adjusted fat intake measured by questionnaire and total energy intake measured by diet record in a validation study (presumably a very good measure of intake). The minimal correlation observed ($r = 0.01$) indicated that the association observed for calorie-adjusted fat intake was not materially confounded by total energy intake. The control of confounding obtained, however, by calorie-adjustment was not complete for all nutrients, further indicating the usefulness of the validation study data as the degree and direction of confounding due to imperfectly measured total energy intake would have been difficult to predict.

SUMMARY

Associations between intake of specific nutrients and disease cannot be considered primary effects of diet if they are simply the result of differences between cases and noncases with respect to body size, physical activity, and metabolic efficiency. In most instances epidemiologic studies of diet and disease should, therefore, be principally directed at the effect of nutrient intakes independent of total caloric intake. This is not always accomplished with nutrient density measures of dietary intake but can be achieved by other methods, including the use of nutrient intakes adjusted for energy intake by regression analysis.

Although pitfalls in the manipulation and interpretation of energy intake data in epidemiologic studies have been emphasized, these considerations also highlight the importance of obtaining a measurement of total energy intake. For instance, if a questionnaire obtained information on only saturated fat intake in a study of coronary heart disease, it is possible that an inverse or no association would be found even if high saturated fat composition of the diet truly caused coronary disease, as the energy intake of cases is likely to be less than that of noncases. Such a finding could be appropriately interpreted if an estimate of total energy intake were available.

The relationships between dietary factors and disease are complex. Even with carefully collected measures of intake, consideration of the biologic implications of various analytic approaches is needed to avoid misleading conclusions.

REFERENCES

Abraham, S. and M. D. Carroll (1979). In Havlik, R. J., and M. Feinlieb, eds.: Food consumption patterns in the United States and their potential impact on the decline in coronary heart disease mortality. Proceedings of the Conference on the Decline in Coronary Heart Disease Mortality. (DHEW publication no. (NIH) 79-1610.

Beaton, G. H., J. Milner, P. Corey, et al. (1979). Sources of variance in 24-h dietary recall data: Implications for nutrition study design and interpretation. *Am. J. Clin. Nutr. 32,* 2546–2559.

Donato, K. and D. M. Hegsted (1985). Efficiency of utilization of various sources of energy for growth. *Proc. Natl. Acad. Sci. (USA) 82,* 4866–4870.

Garabrant, D. H., J. M. Peters, T. M. Mack, and L. Bernstein (1984). Job activity and colon-cancer risk. *Am. J. Epidemiol. 119,* 1005–1014.

Garcia-Palmieri, M. R., P. Sorlie, and J. Tillotson et al. (1980). Relationship of dietary intake to subsequent coronary heart disease incidence: The Puerto Rican Heart Health Program. *Am. J. Clin. Nutr. 33,* 1818–1827.

Garrow, J. S. and J. Webster (1985). Are pre-obese people energy thrifty? *Lancet 1,* 670–671.

Gordon, T., A. Kagan, Garcia-Palmieri, et al. (1981). Diet and its relationship to coronary heart disease and death in three populations. *Circulation 63,* 500–515.

Gordon, T., M. Fisher, and B. M. Rifkind (1984). Some difficulties inherent in the inter-

pretation of dietary data from free-living populations. *Am. J. Clin. Nutr. 39,* 152–156.

Greenland, S. (1980). The effect of misclassification in the presence of covariates. *Am. J. Epidemiol. 112,* 564–569.

Hegsted, D. M. (1985). Dietary standard: Dietary planning and nutrition education. *Clin. Nutr. 4,* 159–163.

Himms-Hagen, J. (1984). Thermogenesis in brown adipose tissue as an energy buffer: implications for obesity. *N. Engl. J. Med. 311,* 1549–1558.

Hofstetter, A., Schutz, E. Jequier, and J. Wahrenj (1986). Increased 24-hour energy expenditure in cigarette smokers. *N. Engl. J. Med. 314,* 79–82.

Horton, E. S. (1983). Introduction: An overview of the assessment and regulation of energy balance in humans. *Am. J. Clin. Nutr. 38,* 972–977.

Howe, G. R., A. B. Miller, and M. Jain (1986). Re: "Total energy intake: implications for epidemiologic analyses" (letter). *Am. J. Epidemiol. 124,* 156–157.

Jain, M., G. M. Cook, F. G. Davis, M. G. Grace, G. R. Howe, and A. B. Miller (1980). A case-control study of diet and colorectal cancer. *Int. J. Cancer 26,* 757–768.

James, W.P.T. and P. Trayhurn (1981). Thermogenesis and obesity. *Br. Med. Bull 37,* 43–48.

Jequier, E. and Y. Schutz (1983). Long-term measurements of energy expenditure in humans using a respiration chamber. *Am. J. Clin. Nutr. 38,* 989–998.

Johnson, M. L., B. S. Burke, and J. Mayer (1956). Relative importance of inactivity and overeating in the energy balance of obese high school girls. *Am. J. Clin. Nutr. 4,* 37–44.

Kromhout, D. and C. L. Doulander (1984). Diet, prevalance and 10-year mortality from coronary heart disease in 871 middle-aged men: The Zutphen Study. *Am. J. Epidemiol. 119,* 733–741.

Kushi, C. H., M. el Lozy, F. J. Stare, et al. (1985). Diet and coronary heart disease (correspondence). *N. Engl. J. Med. 313,* 119–120.

Lyons, J. L., J. W. Gardner, D. W. West, et al. (1983). Methodologic issues in epidemiologic studies of diet and cancer. *Cancer Res. 43,* 2392S–2396S.

McGee, D., D. Reed, and K. Yano (1984). The results of logistic analyses when the variables are highly correlated: An empirical example using diet on CHD incidence. *J. Chronic. Dis. 37,* 713–719.

McGee, D. L., D. M. Reed, K. Yano, et al. (1984). Ten-year incidence of coronary heart disease in the Honolulu Heart Program: Relationship to nutrient intake. *Am. J. Epidemiol. 119,* 667–676.

Mellits, E. D. and D. B. Cheek (1970). The assessment of body water and fatness from infancy to adulthood. *Monogr. Soc. Res. Child. Dev. 140,* 12–26.

Miller, D. S. (1973). Overfeeding in man. In: Berg, G. A., ed. *Obesity in Perspective* (DHEW publication no. (NIH) 75–708.

Morris, J. N., J. W. Marr, and O. B. Clayton (1977). Diet and heart: A postscript. *Br. Med. J. 2,* 1307–1314.

National Center for Health Statistics (1979). Weight and height of adults 18–74 years of age: United States, 1971–74. Hyattsville, Md.: National Center for Health Statistics, 1979. (Series 11, no. 211) (DHEW publication no. (PHS) 79-1659.

NIH Consensus Conference (1985). Lowering blood cholesterol to prevent heart disease. *J.A.M.A. 253,* 2080–2086.

Paffenbarger, R. S.: Jr., A. L. Wing, and R. T. Hyde (1978). Physical activity as an index of heart attack risk in college alumni. *Am. J. Epidemiol. 108,* 161–175.

Prentice, A. M., A. E. Black, W. A. Coward, et al (1986). High levels of energy expenditure in obese women. *Br. Med. J. 292,* 983–987.

Ravussin, E., S. Lilloja, T. E. Anderson, L. Christin, and C. Bogardus (1986). Determinants of 24-hour energy expenditure in man: Methods using a respiratory chamber. *J. Clin. Invest. 78,* 1568–1578.

Romieu, I., W. C. Willett, M. J. Stampfer, G. A. Colditz, L. Sampson, B. Rosner, C. H. Hennekens, and F. E. Speizer (1988). Energy intake and other determinants of relative weight. *Am. J. Clin. Nutr. 47,* 406–412.

Shekelle, R. and Z. Nichaman (1987). Re: Total energy intake: implications for epidemiologic analyses (letter). *Am. J. Epidemiol. 126,* 980.

Shekelle, R. B., O. Paul, and J. Stamler (1985). Diet and coronary heart disease (letter). *N. Engl. J. Med. 313,* 120.

Sidney, S., G. C. Friedman, and R. A. Hiatt (1986). Serum cholesterol and large bowel cancer: a case-control study. *Am. J. Epidemiol 124,* 33–38.

Sims, E.A.H., E. Danforth, E. S. Horton, et al. (1973). Endocrine and metabolic effects of experimental obesity in man. *Recent Prog. Hormone Res. 29,* 457–496.

Sjostrom, L. (1985). A review of weight maintenance and weight change in relation to energy metabolism and body composition. Recent advances in obesity research. Vol IV. Proceedings of the 4th international congress on obesity. Westport, Conn.: Food and Nutrition Press.

Sopko, G., D. R. Jacobs, Jr., and J. L. Taylor (1984). Dietary measures of physical activity. *Am. J. Epidemiol. 120,* 900–911.

Thompson, A. M. and W. Z. Billewicz (1961). Height, weight and food intake in men. *Br. J. Nutr. 15,* 241–252.

Todd, K. S., M. Hudas, and D. H. Calloway (1983). Food intake measurements: Problems and approaches. *Am. J. Clin. Nutr. 37,* 139–146.

Van Itallie, T. B. (1978). Dietary fiber and obesity. *Am. J. Clin. Nutr. 31,* 543–552.

Vena, J. E., S. Graham, M. Zielezny, M. K. Swanson, R. E. Barnes, and J. Nolan (1985). Lifetime occupational exercise and colon cancer. *Am. J. Epidemiol. 122,* 357–365.

Webb, P. (1985). The exchange of matter and energy in lean and overweight men and women: A colorimetric study of over eating, balanced intake and undereating. *Int. J. Obesity 9 (s2),* 139–145.

Willett, W. C. (1987). Implications of total energy intake for epidemiologic studies of breast and large bowel cancer. *Am. J. Clin. Nutr. 45,* 54–60.

Willett, W. C., L. Sampson, M. J. Stampfer, et al. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol. 122,* 51–65.

Willett, W. and M. J. Stampfer (1986). Total energy intake: Implications for epidemiologic analyses. *Am. J. Epidemiol. 124,* 17–27.

Woo, R., R. Daniels-Kugh, and E. S. Horton (1985). Regulation of energy balance. *Ann. Rev. Nutr. 5,* 411–433.

12

Correction for the Effects of Measurement Error

All biologic and physical measurements in any branch of science have error; to a large extent, increments in knowledge depend on reducing this inexactness. It is, therefore, critical to improve continually the technical aspects of exposure measurement, whether based on questionnaires, biochemical assays, or anthropometry. At some level, however, it is difficult or impractical to reduce measurement error further. It is then important to measure the magnitude of the error and evaluate its effect on relationships under investigation. If the effect of measurement error is appreciable, then it may be appropriate to consider a statistical correction to better approximate the relationship that would have been observed if no measurement error had been present.

The correction of estimates of association for measurement error has rarely been employed in the epidemiologic literature. The topic, however, is currently receiving considerable attention and it is likely to become a frequent procedure. Because the methods for correction depend on the form of error, different types of errors and their impact on epidemiologic measures of associations are discussed before considering alternatives for statistical correction. Additional reviews of the statistical effects of measurement errors and methods to compensate for distortion have been published (Snedecor, 1968; Byars and Gail, in press).

TYPES OF ERRORS

The specific sources of error are innumerable; however, they can be thought of as two general types: random and systematic. For random error, the average value of many repeated measures approaches the true value, that is, the law of large numbers applies. For systematic errors, the mean of repeated measurements does not approach the true value. In epidemiologic studies, random or systematic errors, or both, can occur at two different levels: within a person, and

between persons. Thus at least four types of error can exist; these are depicted in Figure 12–1.

Random within-person error is typified by the day-to-day fluctuation in dietary intake discussed in Chapter 3. This apparently random variation is due both to the changes in food intake from day-to-day as well as to errors in the measurement of intake on any one day. Although it may be argued that true variation over time is not really error, it may be considered so if the long-term average intake for an individual is conceptually the true intake for that subject. The distinction between random measurement error and true random day-to-day change in diet is thus usually not important when considering their effects on epidemiologic associations.

In addition to random error, repeated measurements of diet within a subject may also be subject to systematic error. This can occur for many reasons; in open-ended methods, such as 24-hour recalls, persons may consciously or unconsciously tend to either deny or exaggerate their food intake. Systematic within-person error is particularly likely to occur when standardized questionnaires are used: an important food item for a subject (but not necessarily for all subjects) may have been omitted from a questionnaire or misinterpreted by a subject. If such a questionnaire is repeated, this same error is likely to recur; thus the mean of many replicate measurements for an individual will not approach that person's true mean.

Most literature addressing the issue of measurement error in nutritional epidemiology (and other fields) is based on the assumption that within-person error

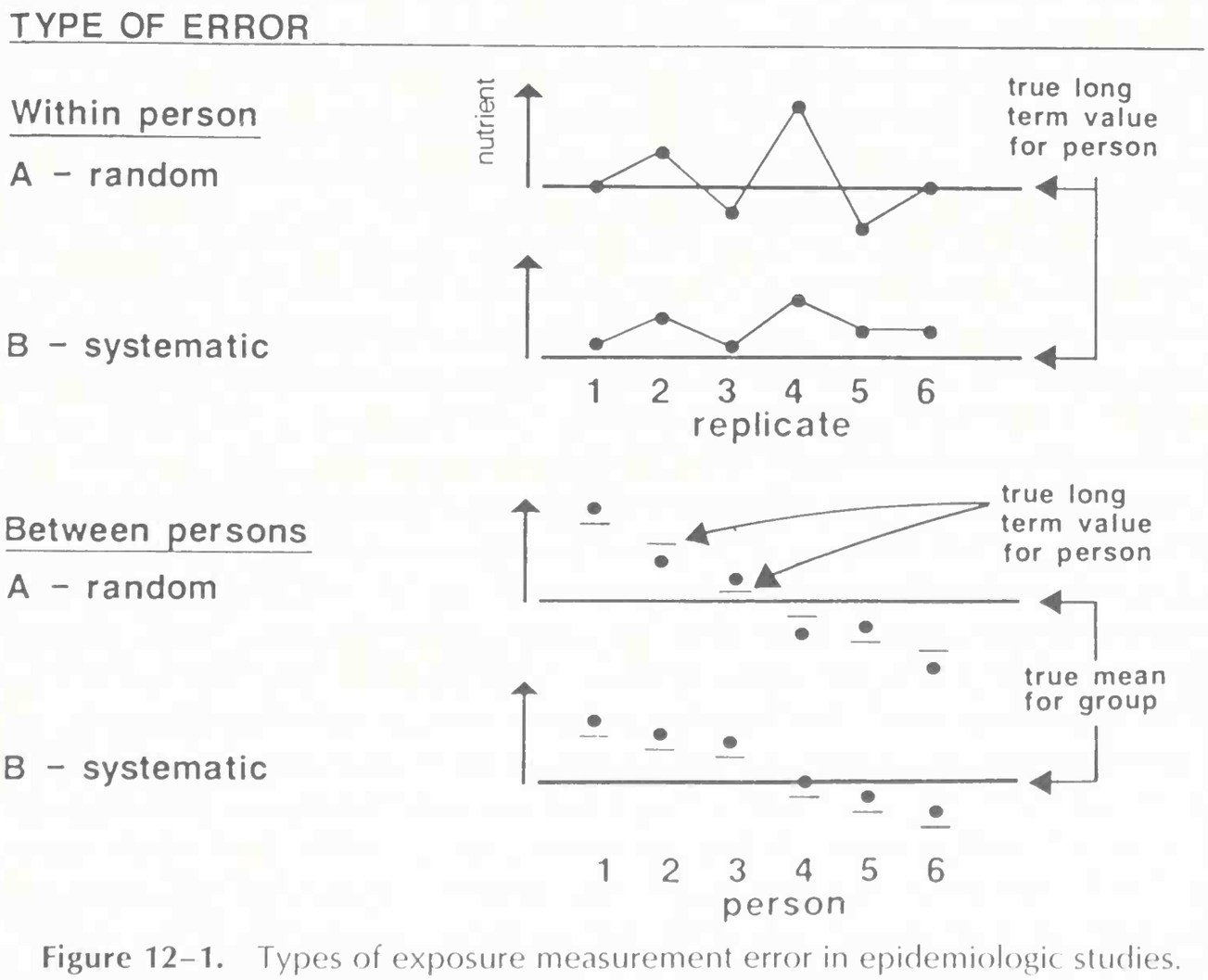


Figure 12–1. Types of exposure measurement error in epidemiologic studies.

is strictly random. This is probably due to two principal reasons: (1) much of statistical theory is based on the assumption of random error and (2) systematic error is considerably more difficult to measure. Random within-person error can be measured simply with a single replicate measure for a sample of subjects, that is, a reproducibility study. The measurement of systematic error requires a second, superior measure of exposure, that is, a validation study. Unfortunately, no perfect measure of true long-term dietary intake exists, and the best measurements (e.g., diet recording or direct observation for many days) are laborious and expensive. Lack of a perfect gold standard is not unique to nutritional epidemiology. For example, it is generally assumed that within-person error in blood pressure measurement is purely random. It is likely, however, that systematic within-person error also occurs, as would happen if some individuals usually develop anxiety when their blood pressure is measured or have anatomic abnormalities in their arm that consistently causes an erroneous reading. A more direct measure of true long-term blood pressure would be continuous monitoring with an intraarterial catheter; however, this is obviously not practical in an epidemiologic setting.

When measuring dietary factors or other exposures among a group of persons, errors can also be either random or systematic. Random between-person error can be either the result of using only one or a few replicate measurements per subject in the presence of random within-person error, or the consequence of systematic within-person errors that are randomly distributed among subjects. Random between-person error implies that an overestimation for some individuals is counter-balanced by an underestimation for others so that the mean for a large group of subjects is the true mean for the group. The standard deviation for the group is, however, exaggerated.

Systematic between-person error results from systematic within-person error that affects subjects nonrandomly. The mean value for a group of persons is thus incorrect. If the systematic error applies equally to all subjects and is simply additive, the observed standard deviation for the group is correct. If individuals, however, are affected to various degrees or the error is multiplicative (for example, proportional to an individual's true level), the standard deviation will also be incorrect. Systematic between-person errors are likely to be frequent and can have many causes: the omission of a commonly eaten food on a standardized questionnaire or the use of an incorrect nutrient composition value for a common food will affect all individuals in the same direction, but not to the same degree as the use of these foods will differ among subjects. As in these examples, it is probably uncommon that systematic within-person error affects all individuals equally. More commonly, random and systematic between-person errors are likely to exist in combination.

Because the ultimate focus of epidemiology is on associations with disease, the impact of exposure measurement error on measures of association, such as relative risk, is of greatest importance. As illustrated in Chapter 3, random within-person error tends to decrease correlation and regression coefficients toward zero and bias relative risks toward one. This effect applies to random between-person errors in general, even if it is the consequence of systematic within-person error that is unequally distributed among subjects. Systematic

errors that affect all persons equally, however, do not affect measures of association.

In the later discussion we assume that all errors apply equally to cases and noncases in an epidemiologic study, that is, that errors are random in relation to disease. Systematic differences in measurement error between these two groups, that is, measurement errors that are biased with respect to disease, have serious consequences that are not readily amenable to correction. This topic is discussed in Chapter 1.

CORRECTION OF EPIDEMIOLOGIC MEASURES OF ASSOCIATION FOR MEASUREMENT ERROR

Several methods are presented here for correcting correlation and regression coefficients and relative risks. Derivations of these formulas and those to compute confidence limits are beyond the scope of this book; the reader needs to refer to the original sources. In using these methods, careful attention must be given to assumptions regarding the type of error (whether it is random or systematic) and the distribution of exposure variables, as normality is often assumed.

Much of the literature related to error correction is based on the model

$$z = x + \epsilon \tag{12-1}$$

where x is the true measure of exposure, z is the surrogate measure that contains error, and ϵ is random measurement error (Kupper, 1984). This model thus assumes that error is simply added to (or subtracted from) the true measurement so that the standard deviation and variance of z are greater than that of x. In this model an assessment of the measurement error can be made by comparing these variances. The ratio of variances, the "reliability coefficient," can be calculated as:

$$r_{zx}^2 = s_x^2/s_z^2 \tag{12-2}$$

The square root of the reliability coefficient (r_{zx}), the correlation between the true and the surrogate measures, is thus the square root of the ratio of the variance of the true measure divided by the variance of the surrogate measure.

Unfortunately, there is little reason to believe that the basic model for this approach to error correction is generally true. In the case of a dietary questionnaire, a very restricted list of foods that are commonly eaten by most people could produce a distribution of nutrient values with a standard deviation that was small, and perhaps even smaller than the true distribution. Equation (12–2) would imply that this restricted questionnaire had little or no error, when in fact it could be providing no useful information. On the other hand, it is possible that a very complete and highly discriminating questionnaire could overestimate intake of those with truly high intake and underestimate intake of those with truly low intake. (This could happen, for example, with a food-frequency questionnaire having multiple choice responses if inappropriately high weights were assigned to the highest category and inappropriately low weights were assigned to the low category when computing nutrient intakes.) Equation (12–

2) would suggest that this questionnaire had a high degree of error, even when discrimination among subjects was excellent. This model of measurement error is appropriate in some situations where the source of error is strictly random within persons and additive; however, it is unlikely to apply to many epidemiologic measures, in particular the assessment of diet by questionnaires. An alternate model relating true and surrogate measures using regression analysis that requires fewer assumptions is discussed later.

Correction of Correlation Coefficients

As discussed in Chapter 3, random within-person error in the measurement of one or both variables being compared tends to reduce correlation coefficients toward zero. This issue and approaches for correcting the correlation coefficients based on a partitioning of variance components have been discussed by Liu and colleagues (1978) and Beaton and colleagues (1979). When both variables, x and y, are subject to random within-person variation the relationship between the true correlation (r_t) and the observed correlation (r_o) can be written

$$r_t = r_o\sqrt{(1 + \lambda_x/n_x)(1 + \lambda_y/n_y)} \qquad (12\text{–}3)$$

where $\lambda_x = s_w^2/s_b^2$ for x, that is, the ratio of the within- and between-person variances for x.

$\lambda_y = s_w^2/s_b^2$ for y
n_x is the number of replicates per person for x variable,
n_y is the number of replicates per person for the y variable.

If only one variable is considered to have no within-person variance, this equation reduces to

$$r_t = r_o\sqrt{1 + \lambda_x/n_x} \qquad (12\text{–}4)$$

The within-person and between-person variances are obtained from an ANOVA model as described in Chapter 3. Alternatively, λ_x and λ_y can be obtained from their respective intraclass correlations (r_I, the correlation describing the reproducibility of x or y, which can be thought of as the average correlation between pairs of replicate measurements):

$$\lambda = (1 - r_I)/r_I \qquad (12\text{–}5)$$

It should be noted that this method for error correction assumes that variation in x and y are independent (e.g., that they fluctuate randomly with respect to each other from day to day). This assumption does not appear to be seriously violated in most instances, but deserves consideration in any particular application. An additional term is required if variation in x and y are not independent (Beaton et al., 1979).

Example: In a study of dietary questionnaire validity (Willett et al., 1985), we examined the correlation between a self-administered food-frequency questionnaire and multiple diet records, assuming that the diet records represented true intake. Realizing that the within-person variation in diet is substantial, we col-

lected 28 days per subject sampled over a 1-year period, with the assumption that the average of these days would reasonably represent an individual's true long-term intake. For cholesterol intake, the observed correlation between the questionnaire and the diet record intake was 0.51. We then sampled only 4 of the 28 days of diet records for each subject; the correlation between the questionnaire estimate of cholesterol intake and the mean of these 4 days per subject was only 0.29. Based on these 4 days of dietary cholesterol intake per subject, an analysis of variance indicated that the within-person variance (s_w^2) was 3.43 and the between-person variance (s_b^2) was 0.50.

From Equation (12–4) the true correlation between a single questionnaire administration and true long-term cholesterol intake can be estimated as:

$$r_t = 0.29\sqrt{1 + (3.43/0.50)/4} = 0.48$$

Thus, correcting the observed correlation for the attenuating effect of random within-person error provides a value similar to that obtained with a large number of replicates. The loss of information from using 4 rather than 28 days of dietary data is reflected in a wider confidence interval for the corrected coefficient. In this example, the 95 percent confidence interval for the original correlation of 0.51 was 0.39 to 0.61; for the corrected correlation of 0.48 this interval was 0.20 to 0.68. The calculation of confidence intervals for corrected correlation coefficients is more complex, as they reflect both error in the estimate of the observed correlation (which is a function of the number of subjects) and error in the estimation of the within-person and between-person variances (which is a reflection of the variability and number of replicates). A formula for this calculation is provided elsewhere (Rosner and Willett, 1988). From the example, it can be seen that the use of corrected correlation coefficients can provide a reasonable estimate of the true correlation with only a few replicate measures per subject (a minimum of two measurements for at least a subsample of subjects is necessary for ANOVA).

In making these corrections, it is important to consider carefully the true relationship that is of interest. In this example, the dietary questionnaire was administered twice so that we could have further corrected the observed correlation for within-person variation in the questionnaire using Equation (12–3). This, however, would have provided an estimate of the correlation between an infinite (or large) number of days of dietary record intake per subject with an infinite (or large) number of questionnaire administrations per subject. This is obviously not useful information as we generally have only one, or at most a few, questionnaire administrations per subject.

In designing a study to measure the correlation between two variables where one or both are subject to random within-person variation, the investigator must decide whether to enroll many subjects and collect few replicates per person, or to enroll few subjects and collect many replicates for each person. An objective in designing such a study is to obtain the most precise estimate of the true correlation, that is, to minimize its standard error or confidence interval. The standard error of the corrected correlation (Rosner and Willett, 1988) thus provides a criterion for optimizing the trade-off between number of subjects and number

of replicates per subject given a fixed number of measurements. The optimal solution is not simple as it depends on both the true correlation and on the intraclass correlation for the variable with random error. Table 12–1 provides variances of corrected correlations, assuming a total of 1000 measurements, for various values of the true correlation (r_t) and the intraclass correlation (r_I). It can be seen, except for extreme situations, that the minimum variance occurs with only two replicates per subject. Even when the minimum variance is provided by a larger number of replicates, two replicates provide a variance that is not far from optimal. The intuitive reason for this finding is that as replicate measures are correlated to at least some degree, each additional replicate adds progressively less new information. Although these particular calculations were based on the allocation of 1000 total measurements (i.e., the number of subjects times the number of replicates per subject), the optimal number of replicates per subject was not sensitive to the total number of measurements. Further calculations using replicate samples for only a subset of the population indicated that it was

Table 12–1. Variances of corrected correlation coefficients (r_t) multiplied by a fixed total number of measurements over the entire sample (N = subjects × measurements/subject = 1000 in this example).

		Number of replicate measures per subject (n_y)								
r_t	r_I	2	3	4	5	6	7	8	9	10
0.1	0.1	11.4	12.1	13.1	14.0	15.0	15.9	16.9	17.9	18.8
0.1	0.3	4.3	5.3	6.3	7.2	8.2	9.2	10.2	11.2	12.1
0.1	0.5	3.0	3.9	4.9	5.9	6.9	7.9	8.8	9.8	10.8
0.1	0.7	2.4	3.4	4.3	5.3	6.3	7.3	8.3	9.2	10.2
0.1	0.9	2.1	3.1	4.0	5.0	6.0	7.0	8.0	8.9	9.9
0.3	0.1	14.3	13.3	13.5	14.0	14.7	15.4	16.1	16.9	17.7
0.3	0.3	4.2	4.9	5.7	6.5	7.3	8.2	9.0	9.8	10.6
0.3	0.5	2.7	3.5	4.3	5.1	6.0	6.8	7.6	8.5	9.3
0.3	0.7	2.1	2.9	3.7	4.6	5.4	6.2	7.1	7.9	8.7
0.3	0.9	1.8	2.6	3.4	4.3	5.1	5.9	6.7	7.6	8.4
0.5	0.1	20.2	15.6	14.5	14.2	14.2	14.4	14.7	15.1	15.5
0.5	0.3	4.1	4.3	4.7	5.2	5.7	6.3	6.8	7.4	7.9
0.5	0.5	2.2	2.7	3.2	3.8	4.3	4.9	5.5	6.0	6.6
0.5	0.7	1.6	2.1	2.7	3.2	3.8	4.3	4.9	5.5	6.0
0.5	0.9	1.2	1.8	2.4	2.9	3.5	4.0	4.6	5.2	5.7
0.7	0.1	29.0	19.2	16.0	14.5	13.7	13.2	12.9	12.8	12.7
0.7	0.3	3.9	3.5	3.5	3.6	3.8	4.0	4.2	4.4	4.6
0.7	0.5	1.6	1.7	1.9	2.2	2.4	2.7	2.9	3.2	3.4
0.7	0.7	0.9	1.1	1.4	1.6	1.9	2.2	2.4	2.7	2.9
0.7	0.9	0.6	0.9	1.1	1.4	1.6	1.9	2.2	2.4	2.7
0.9	0.1	40.9	24.1	18.3	15.3	13.5	12.2	11.2	10.5	9.9
0.9	0.3	3.9	2.7	2.2	2.0	1.9	1.8	1.8	1.7	1.7
0.9	0.5	1.0	0.8	0.8	0.7	0.7	0.7	0.8	0.8	0.8
0.9	0.7	0.3	0.3	0.3	0.4	0.4	0.4	0.5	0.5	0.5
0.9	0.9	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4

Variances (× 1000) are shown for a given number of replicates per subject (n_y), a given intraclass correlation (r_I), and the corrected correlation coefficient (r_t).

From Rosner and Willett, 1988.

generally most efficient (in minimizing the standard error of the corrected correlation coefficient) to have at least two samples for all subjects; the exception occurred when the intraclass correlation was very high. In addition to statistical efficiency, practical considerations also influence the optimal allocation of measurements. Applications of this approach have important implications for the design of validation studies; these are discussed in Chapter 6.

Systematic within-person error, as would typically characterize standardized dietary questionnaires, also tends to attenuate correlation coefficients toward zero as this error may not affect all persons equally. The methods used to correct for random within-person error, however, are not applicable, and it is not apparent that this issue has been formally addressed. Systematic between-person errors in which all subjects are measured too high or too low by the same amount do not affect the correlation coefficient, thus no correction is needed.

Correction of Regression Coefficients

The correction of regression coefficients for random within-person error has been discussed by Beaton and colleagues (1979), Liu and colleagues (1978), Madansky (1959), and others (Tukey 1951; Riggs et al., 1978). When random error affects only the dependent variable (y), the regression coefficient is not attenuated and no correction is needed. When such error affects the independent variable (x), a simple correction can be made as described by Beaton and coworkers (1979):

$$b_t = b_o(1 + \lambda_x/n_x) \tag{12–6}$$

where b_t = the true regression coefficient, b_o = the observed regression coefficient, and the other notation is the same as for Equation (12–4). Again, this assumes that distributions are reasonably normal and that the within-person variation is random. Due to the measurement error, the true confidence interval is wider than the observed interval; a standard method does not appear to have been published for calculating the corrected values.

Example: Several authors have used the relationship of dietary cholesterol intake and serum cholesterol level to illustrate the attenuating effects of random within-person variation on regression coefficients (Beaton et al., 1979; Liu et al., 1978). Based on metabolic ward feeding studies, Mattson and co-workers (1972) observed that serum cholesterol increased by 12 mg/dl for each increase of 100 mg of cholesterol intake per 1000 kcal daily.Thus b_t in this example is 0.12. From Table 3–3, it can be seen that the value for the ratio of within-person to between-person variances (λ) for cholesterol intake (divided by calories) is approximately 5 to 13. Assuming a value of 5 and that only 1 day of observation is obtained, the observed regression coefficient would be

$$\begin{aligned} b_o &= b_t/(1 + \lambda_x/n_x) \\ &= 0.12/(1 + 5/1) = 0.02 \end{aligned}$$

Because the attenuation is large and many factors other than dietary cholesterol intake affect serum cholesterol levels (e.g., laboratory error, genetic factors, phys-

ical activity, other dietary factors), it is not surprising that most studies based on a single or a few days of intake have failed to observe any association. Moreover, the data of Mattson and co-workers demonstrate that serum cholesterol is only weakly influenced by diet within the range that they studied; because a typical cholesterol intake would be 150 mg/1000 calories and a typical serum level would be 250 mg/dl, a doubling of intake (to 300 mg/1000 calories) would result in less than an 8 percent rise in serum cholesterol (18 mg/dl). For this reason, serum cholesterol cannot be considered a useful indicator of cholesterol intake or a reasonable standard for validating any method of measuring dietary cholesterol intake.

Analogous to the correction of correlation coefficients, methods for correcting regression coefficients based on the ratio of within-person to between-person variances are not applicable when systematic within-person error is present. In this case, a separate validation study employing an independent, more accurate measure of true exposure is necessary. A two-step process can be used. First, the "true" measure (x) is regressed on the surrogate measure (z) to obtain a regression coefficient (γ).

$$x = \alpha + \gamma z + \epsilon \tag{12–7}$$

This coefficient (γ) thus describes the true change in exposure that corresponds to a given observed change in the surrogate exposure. As a second step, this information is used to correct the observed regression coefficient (b_o) describing the relationship between the surrogate exposure (z) and the dependent variable ($y = \alpha + b_o z + \epsilon$). This relationship between the true and observed values of the regression coefficient is simply

$$b_t = b_o/\gamma \tag{12–8}$$

Example: Using the relationship between dietary cholesterol and serum cholesterol described previously, let us assume, in a hypothetical cross-sectional study, that the observed relationship between cholesterol intake (in mg/1000 kcal daily) measured by a dietary questionnaire and serum cholesterol (mg/dl) was 0.06. In a separate validation study comparing the questionnaire with a "true" measure of dietary intake (such as a large number of days of diet recording per subject), a 100-mg/1000 kcal daily change in cholesterol intake measured by the questionnaire corresponded to a 50-mg/1000 kcal daily change according to the true method, that is, $\gamma = 0.5$. The corrected regression coefficient, would thus be $b_t = 0.06/0.5 = 0.12$

Correction of Relative Risks

Random and systematic within-person errors that are unrelated to disease status tend to bias relative risks toward the null value of one (see Chapter 3). Several methods are available to correct these attenuated relative risk estimates; methods are discussed first for categorical, then for continuous variables.

Categorical Variables

This discussion on measurement error has so far only considered continuous variables as most dietary exposures are of this form. When computing relative risks, exposures may still be considered as continuous variables; the rates for two points on the distribution of exposure can be compared using a statistical model, such as a logistic or proportional hazards model. More commonly, the continuous exposure is broken into categories, such as quintiles or 100-g units, and relative risks are computed comparing these groups. The following example illustrates attentuation of the relative risk calculated as an odds ratio from a 2 $\times$ 2 table.

Example: Let the true relation between exposure and disease defined by a 2 $\times$ 2 table be as follows:

		Exposed (E)	Unexposed (U)	Total (E + U)
Diseased	(D)	a = 400	b = 600	1000
Well	(W)	c = 200	d = 800	1000

The odds ratio (OR) = $(a \div b)/(c \div d)$ = (400 ÷ 600)/(200 ÷ 800) = 2.67.

If the true relationship is studied with an imperfect measurement method, the performance of the surrogate method can be described in terms of ϕ = sensitivity (the proportion of truly exposed subjects who are identified as exposed by the epidemiologic method) and ψ = specificity (the proportion of subjects who are truly unexposed who are identified as unexposed by the method). For this example, assume that an exposure measurement method having a sensitivity of 0.6 and a specifity of 0.9 is used to study the true exposure-disease relationship described previously. Therefore, of the 400 diseased and truly exposed subjects, only 400 $\times$ 0.6 = 240 would be measured as such and 400 − 240 = 160 would be measured as unexposed. Of the 600 diseased and truly unexposed subjects, only 600 $\times$ 0.9 = 540 would be measured as unexposed and 600 − 540 = 60 would be called exposed. Thus 240 + 60 = 300 diseased subjects would be measured as exposed and 540 + 160 = 700 as unexposed. With similar calculations for the well subjects, the observed 2 $\times$ 2 table would thus be:

	Exposed (b)	Unexposed (U)		Exposed (E)	Unexposed (U)
Diseased	240 + 60	540 + 160	Diseased	300	700
Well	120 + 80	720 + 80	Well	200	800

observed OR = (300 ÷ 700)/(200 ÷ 800) = 1.71

In this example the observed odds ratio (1.71) is appreciably lower than the true odds ratio (2.67) due to measurement error.

Barron (1977) has provided a method of relative risk correction suitable for a simple 2×2 table when errors exist in the measurement of either (or both) exposure and disease. This method makes no assumptions regarding the nature of the measurement error, but does require an external validation study based on a "true" measure of exposure to define the probabilities of misclassification. The results of this external validation can be expressed in terms of sensitivity (ϕ) and specificity (ψ), as described previously. Barron has described his method in terms of matrix algebra, which includes the possibility of error in the diagnosis of disease as well as exposure. Here it will be assumed that no error exists in the diagnosis of disease (this is often reasonably true in many epidemiologic studies) and the method of correction is presented in simple algebraic form. The use of this method is discussed in more detail by Kleinbaum and colleagues (1982).

The true relationship between exposure and disease can be described for a 2×2 table as using the notation:

	Exposed	Unexposed
Diseased	a^*	b^*
Well	c^*	d^*

Correspondingly, the observed relationship using an imperfect exposure measurement method can be expressed as

	Exposed	Unexposed	Total
Diseased	a	b	$a + b = n_D$
Well	c	c	$c + d = n_W$

Given the sensitivity (ϕ) and the specificity (ψ) of the exposure measurement, the true values for a^*, b^*, c^*, and d^* corrected for measurement error can be written:

$$a^* = (n_d\psi - b)/(\phi + \psi - 1)$$
$$b^* = (n_d\phi - a)/(\phi + \psi - 1)$$
$$c^* = (n_w\psi - d)/(\phi + \psi - 1)$$
$$d^* = (n_w\phi - c)/(\phi + \psi - 1)$$

Example: Referring back to our previous example that demonstrated how misclassification attenuated the correct relative risk, this method can be used to reconstruct the true relationship. If, from a separate validation study, we know that $\phi = 0.6$ and $\psi = 0.9$, then:

$$a^* = (1000 \times 0.9 - 700)/(0.6 + 0.9 - 1) = 400$$
$$b^* = (1000 \times 0.6 - 300)/(0.6 + 0.9 - 1) = 600$$
$$c^* = (1000 \times 0.9 - 800)/(0.6 + 0.9 - 1) = 200$$
$$d^* = (1000 \times 0.6 - 200)/(0.6 + 0.9 - 1) = 800$$

Using these corrected cell numbers, the odds ratio = $400 \div 600/200 \div 800 = 2.67$, the same value as obtained for the original 2×2 table.

Although this method is straightforward, and provides a simple estimate of the effects of misclassification, it has several important limitations. First, in practical epidemiologic applications we are seldomly interested in a single 2×2 table; typically age, gender, and other risk factors need to be controlled in any analysis. Furthermore, a method to compute corresponding confidence limits has not been provided. (It might well be reasonable to compute a test of significance based on the observed data, and then use the test-based method described by Miettinen (1976) in combination with the corrected point estimate for the relative risk; however, this would not account for error in the estimation of the sensitivity and specificity. This approach has apparently not been formally explored.)

Continuous Variables

Determination of True Standard Deviations. A simple approach can be based on an extension of Figure 3–5, which depicts how random within-person error reduces the true magnitude of association.

Using replicate measurements from all or a sample of study subjects, analysis of variance can be employed to separate the within- and between-person components of variation (s_w^2 and s_b^2). The square root of the between-person variance (s_b) is the true standard deviation that would have been observed if no within-person variation had been present, that is, if we had used the mean of a large number of replicates per subject. Because the mean for cases and the mean for controls are not biased by this type of error, the true means and standard deviations for cases and controls can be described by the observed means $\pm$ s_b. If subjects above a cutpoint (x) are considered exposed, then we can calculate the proportion of cases and controls above any specific cut-point using a table of normal deviates. These proportions can then be used to calculate the true relative risk or odds ratio.

Example: The data employed to generate Figure 3–6B can be used in reverse to obtain a corrected relative risk. In this example, normal distributions for intake of a nutrient were observed for a case and a noncase group with the mean of the cases being 0.25 observed standard deviations above the mean of the noncases. A cut-point 0.5 observed standard deviations above the mean of the noncase group was used to define exposed (above this point) and unexposed (below this point) subjects. Using a table of normal deviates to obtain the areas above and below the cut-point for cases and noncases, the proportions of exposed cases and noncases are:

	Nutrient Intake	
	Exposed	Unexposed
Cases	0.41	0.59
Noncases	0.31	0.69
	OR = 1.55	

Now, let us assume that the error in the measure of exposure is strictly random within-person error and that a reproducibility study has been conducted in which it was determined that the within-person variance was three times the true between-person variance ($s_w^2/s_b^2 = 3$). Thus, the total observed variance is $s^2_{obs} = s_b^2 + 3s_b^2 = 4s_b^2$. Therefore, $s_{obs}/2 = s_b$. That is, the true between-person standard deviation (s_b) is only one-half the observed standard deviation. Therefore, the same cut-point actually corresponds to 1.0 standard deviation units above the noncase mean. Returning to the table of normal deviates to determine the areas above and below that cut-point, the following proportions are obtained:

	Nutrient Intake	
	Exposed	Unexposed
Cases	0.31	0.69
Noncases	0.16	0.84
	OR = 2.36	

The odds ratio obtained after removing the effect of within-person variation is thus deattenuated and corresponds to the true value determined in Figure 3–5A.

This approach is intuitively simple and provides a conceptual cornerstone, but has several limitations including the requirement for normally distributed variables (which may usually be approximated by transformation), the present lack of a method for obtaining confidence limits, and the inability to adjust simultaneously for confounding factors.

More complex approaches for correcting relative risk estimates obtained from multiple logistic regression models for random within-person exposure error have been provided by Carroll and colleagues (1984), Stefanski and Carroll (1985), Armstrong (1985), and Kaldor and Clayton (1985). Prentice (1982) has published a method of correction for use in proportional hazards models. A relatively simple method for providing approximate corrections for measurement error, whether due to random or systematic within-person error is described later.

The single imputation method approach provides an approximate correction for errors in continuous exposure variables. First, a validation or calibration study conducted among a subset of participants or similar subjects is used to correct the observed exposure value for each individual using regression analysis. Briefly, for individuals in the validation study, a simple linear regression model is fitted with x equal to the true value and z equal to the observed value [see Equation (12–7)]. The "corrected value" for each individual is their predicted value based on this model. These corrected values are then used in further analysis, whether they be stratified or multivariate analyses. It is apparent that relative risks based on ranked variable categories (e.g., percentile or quintiles) after the correction has been made are not corrected by this method, as the relative rankings of individuals are unchanged. Relative risks based on absolute exposures, however, are corrected. For example, if we are interested in the rel-

ative risk of colon cancer for a 20-g difference in daily fat intake, this method would be appropriate. This approach is further limited because it does not provide appropriate confidence limits as error in the estimation of corrected values is not incorporated.

The linear approximation method of Rosner and colleagues (1989) extends the simple imputation method for relative risk correction to multiple logistic regression. This approach accounts for both random and systematic within-person measurement error. Furthermore, the method provides tests and confidence limits that incorporate uncertainty due to both sample size in the main study as well as that due to uncertainty in the estimation of validity. The latter component is important, because when validation studies are small, the degree of correction applied to the observed relative risk itself has error. Because this method is based on a multiple logistic regression model, it can provide adjustment of the observed relative risk for the effect of other variables before correcting for measurement error. In common with the other methods to correct for systematic within-person error, this method requires data from a separate validation study. Because of the relative simplicity of this method and its ready applicability to epidemiologic data, it is discussed in detail.

The logistic model relating disease and observed exposure (z) can be written

$$\ln\{p/(1 - p)\} = \alpha + \beta z \tag{12–9}$$

while the true relationship between disease and exposure can be written

$$\ln\{p/(1 - p)\} = \alpha^* + \beta^* x \tag{12–10}$$

Using data from the separate validation study, the relationship between the observed exposure (z) and true exposure (x), as in Equation (12–7), can be written

$$x = \alpha' + \gamma z + \epsilon \tag{12–11}$$

From substitution of Equation (12–11) into Equation (12–10)

$$\beta^* = \beta/\gamma \tag{12–12}$$

Because antilog of the logistic regression coefficient is an estimate of the relative risk, the corrected relative risk (RR_t) is simply computed as

$$RR_t = \exp(\beta/\gamma) \tag{12–13}$$

It is of note that the corrected value for β^* in Equation (12–11) is an approximation rather than an exact equality, as the logistic model is not linear. For this reason the correction is not complete and some bias toward the null remains. However, with small or modest relative risks (less than about five) and a reasonably stable estimate of γ, the degree of residual bias is generally small and unimportant (Rosner et al., 1989).

A formula for the variance of β^*, which incorporates error in β as well as error in the estimation in validity (γ) can be written

$$\text{var}(\beta^*) = (1/\gamma^2)(\text{var}\beta) + (\beta^2/\gamma^4)(\text{var}\gamma) \tag{12–14}$$

where varβ is obtained from Equation (12–9) and varγ from Equation (12–11).

It follows that the 100 percent $\times$ $(1 - \alpha)$ confidence limits are given by

$$\exp[\beta^* \pm Z_{1-\alpha/2}\text{s.e.}(\beta^*)] \qquad (12\text{–}15)$$

In some instances, the estimate (x) of true exposure may be subject to random within-person error, such as when the average of a small number of replicate measures is used to estimate true exposure. Because the estimate of true exposure (x) is used in Equation (12–11) as the dependent variable, γ still provides an unbiased estimate of the regression coefficient of x on z when x is subject to random within-person variation. Because it appears that within-person errors based on diet records can be reasonably assumed to be random in some instances, the validation study could be conducted with as little as a single day of diet recording per subject. Random error inherent in x, however, results in increased variance of γ and, ultimately, in a wider confidence interval for the corrected relative risk.

Example: Relative risk of breast cancer in relation to calorie-adjusted saturated fat intake corrected for errors in measurement (continuous variable).
In a previous study, we examined prospectively the relationship between dietary saturated fat intake and risk of breast cancer among a cohort of 89,538 women (Willett et al., 1987b). To measure dietary fat in this study, we used a self-administered, semiquantitative food-frequency questionnaire that had been subjected to a detailed validation study (Willett et al., 1985) among 173 cohort members. To represent true dietary intake in the validation study, we used an average of four 1-week diet records based on weighed food intake collected by each subject at 3-month intervals over a 1-year period.

In the main study, we employed a multiple logistic model with breast cancer incidence as the dependent variable and calorie-adjusted saturated fat intake (a continuous variable) as the primary predictor variable. Age and alcohol intake were also included as covariates. In this model, the coefficient for calorie-adjusted saturated fat intake represents the effect on breast cancer incidence of a daily increase of 1 g of saturated fat with total calories held constant. Because the average saturated fat intake was approximately 25 g daily, 1 g is a very small increment. Thus, we computed the relative risk for a 10-g increase in daily calorie-adjusted saturated fat intake, which approximates the difference between the means of the top and bottom quintiles as assessed by the diet record data (Willett et al., 1985). To estimate γ, we used the data provided by the 173 women in the validation study: calorie-adjusted saturated fat intake measured by the average of four 1-week diet records (x) was regressed on the calorie-adjusted saturated fat intake measured by the questionnaire (z). The regression coefficient for the questionnaire measurement thus provides an estimate of γ, and var(γ) is obtained directly from this model; these two parameters were then used in Equations (12–13) and (12–14).

Based on the full 28 days of diet records, the estimated slope (γ) representing the change in diet record calorie-adjusted saturated fat intake (grams/day) associated with a 1-g change in daily saturated fat intake measured by the questionnaire was 0.468, with a standard error of 0.048. Based on the actual data in the

total cohort, the observed relative risk for a 10-g increase in calorie-adjusted saturated fat intake was 0.92 (95% CI = 0.80 to 1.05). Using Equation (12–13), the corrected relative risk was 0.83 with 95 percent confidence limits from 0.61 to 1.12 (Table 12–2), providing the best estimate of the effect due to a true 10-g difference in calorie-adjusted saturated fat intake given our main study data and allowing for measurement error. As expected, the point estimate of the corrected relative risk is further from unity than the observed relative risk, and the width of the 95 percent confidence limits has increased. For purposes of illustration, hypothetical examples are provided using different values for β but the same standard error of β and the same value for γ (Table 12–2).

In this example, we assumed that the 28 days of diet recording in the validation study were sufficient to dampen within-person variation and thus, provide a close approximation to true intake for an individual. To provide an example of the situation where x (true exposure) is still subject to within-person variation, we sampled 2 days (one each from weeks 1 and 3) and 4 days (one from each week) from the total 28 days of diet record. In Table 12–2, we have provided corrected relative risk estimates using these 2 or 4 days of diet recording to estimate γ. As expected, the estimates of γ were similar, but their standard errors were larger; in this instance, however, the degree of variation in the standard error of γ had only a small influence on the confidence limits for the corrected relative risks. Because, by chance, the value for γ was somewhat lower using all 28 days, the confidence intervals were actually slightly narrower with fewer days. The application of this method of correction for exposure data based on ordinal categories (e.g., quintiles) is discussed in Rosner and colleagues (1989).

Table 12–2. Relative risks of breast cancer for a 10-g/day increase in calorie-adjusted saturated fat intake, adjusted for errors in measurement

Example	Validation study[a]			Main study[b]: RR for 10-g increase			
				Observed		Corrected	
	Dietary questionnaire vs	γ	SE	RR	(95% CI)	RR	(95% CI)
1. Actual data	28 day record	0.468	(0.048)	0.92	(0.80, 1.05)	0.83	(0.61, 1.12)
$\beta = -0.0878$	4 day record	0.554	(0.073)	0.92	(0.80, 1.05)	0.85	(0.66, 1.10)
SE = 0.0712	2 day record	0.540	(0.090)	0.92	(0.80, 1.05)	0.85	(0.65, 1.11)
2. Hypothetical data	28 day record	0468	(0.048)	1.00	(0.87, 1.15)	1.00	(0.74, 1.35)
$\beta = 0.000$	4 day record	0.554	(0.073)	1.00	(0.87, 1.15)	1.00	(0.78, 1.29)
SE = 0.0712	2 day record	0.540	(0.090)	1.00	(0.87, 1.15)	1.00	(0.77, 1.29)
3. Hypothetical data	28 day record	0.468	(0.048)	1.50	(1.30, 1.72)	2.38	(1.68, 3.36)
$\beta = 0.405$	4 day record	0.554	(0.073)	1.50	(1.30, 1.72)	2.08	(1.52, 2.85)
SE = 0.0712	2 day record	0.540	(0.090)	1.50	(1.30, 1.72)	2.12	(1.48, 3.02)

[a]Validation study data are based on 173 women participating in the Nurses' Health Study. Each woman completed four 1-week diet records over a 1-year period. The 4 days used in this analysis were sampled one from each week; the 2 days were sampled one from each of weeks 1 and 3.

[b]Main study data are based on 590 cases of breast cancer occurring among 89,538 women participating in the Nurses' Health Study, aged 34–59 years and followed for 4 years. Logistic model included calorie-adjusted saturated fat intake in g/day as a continuous variable, age (34–39/40–49/50–54/55–59) and alcohol intake (0, 0.1–4.9, and 5+ g/day).

From Rosner et al. (1989).

Ideally, the validation study would be conducted among a subsample of the main study participants to minimize concern regarding the generalizability of the validity estimate to the main study population. It may be possible, with caution, however, to use an estimate of validity from a completely external source, if the populations are generally similar.

The situation in which measurement error is strictly due to random within-person variation (and no "standardization" has been employed) can be considered as a special case of the method of relative risk correction indicated in Equation (12–12). In this situation, the validation study can consist of one or more repeated measures of exposure among a subsample of subjects. If a large number of replicate measures are obtained per subject, the mean of these can be taken as an approximation of the true measure (x) and γ can be obtained from the regression coefficient of the mean exposure on the observed exposure using a single measure in Equation (12–11). Alternatively, γ can be estimated by the intraclass correlation between replicates for persons in the validation study (Donner and Koval, 1980).

Equation (12–13) may be rearranged to compute the relative risk that would be observed given an estimated true relative risk and a particular level of validity (γ):

$$RR_o = (RR_t)^{\gamma} \tag{12–16}$$

where RR_o is the observed relative risk and RR_t is the estimated true relative risk. In Table 12–3, examples of observed relative risks corresponding to estimated true relative risks of 1.5, 2.0, 3.0, and 5.0, and values of γ ranging from 0.2 to 1.0 are provided. There is no obvious threshold of γ below which a measure of exposure is useless; however, for true relative risks of 1.5 or 2.0, epidemiologic effects are difficult to detect if γ is substantially less than 0.5.

It will be apparent that the value of γ does not readily provide a generally interpretable measure of validity. For example, in Table 12–2, part of the cor-

Table 12–3. Observed relative risks for different levels of validity in the measurement of exposure[a]

	True relative risks			
γ[b]	1.5	2.0	3.0	5.0
0.2	1.08	1.15	1.25	1.38
0.3	1.13	1.23	1.39	1.62
0.4	1.18	1.32	1.55	1.90
0.5	1.22	1.41	1.73	2.24
0.6	1.28	1.52	1.93	2.63
0.7	1.33	1.62	2.16	3.09
0.8	1.38	1.74	2.41	3.62
0.9	1.44	1.87	2.69	4.26
1.0	1.50	2.00	3.00	5.00

[a] $RR_o = (RR_t)^{\gamma}$ where RR_o is the observed relative risk and RR_t is the estimated true relative risk.

[b] The regression coefficient for the true measure on the surrogate measure or (when both measures have the same standard deviation) the correlation coefficient between them.

rection is related to a change in scale as the standard deviation of calorie-adjusted saturated fat intake measured by the questionnaire is larger than the standard deviation according to the diet record. Thus, even if the two measures were perfectly correlated, the larger standard deviation of the questionnaire would result in a value of γ less than one. Although imperfect measures usually result in larger variances, this is not necessarily true, so that the value of γ can exceed one (this can also result simply from a change in units). If both the "true" (x) and surrogate (z) measures have the same standard deviation, then the value for γ is equivalent to the correlation coefficient relating x and z, and is comparable to the measure of misclassification discussed by Walker and Blettner (1985).

The effect of including covariates in a multiple logistic model on the correction of relative risks deserves comment. In principle, the measure of validity should be conditional on the same set of covariates that would be included in the logistic model; this would be particularly important for variables that are strongly associated with the primary exposure. Thus, both x and z should each be adjusted for relevant covariates before regressing x on z as in Equation (12–11) (this adjustment can also be accomplished by including these covariates as independent variables). For example, if the multiple logistic model in a nutritional analysis included gender, then the estimate of validity should be conditional on gender as men tend to eat more of most nutrients than women and we have shown that adjustment for gender reduces the apparent association between the questionnaire measurement of intake and a diet record measurement (Willett et al., 1987a).

ESTIMATION OF RELATIVE RISKS BASED ON DUAL RESPONSES

Marshall and Graham (1984) have described an approach to improve the estimation of relative risks based on two measures of exposure for all subjects, with the assumption that neither method is perfect. For dichotomous variables they suggest that subjects be categorized as either exposed according to both measures, unexposed according to both measures, or exposed according to only one measure. The true relative risk would then be best approximated by comparing subjects classified as exposed by both methods with those classified as unexposed by both methods. This approach provides an opportunity to improve the estimation of effects in epidemiologic studies in the absence of known validity for either measurement. This lack of quantified measurement error, however, limits interpretation as relative risks still tend to be underestimated, although the magnitude of underestimation remains uncertain. Walter (1984) proposed a maximum likelihood method based on dual responses to provide an unbiased estimate of the true association between exposure and disease. This method is also more efficient because all the data are used, not just those with concordant exposures. Walter noted that both the maximum likelihood approach and the method suggested by Marshall and Graham assume that errors are uncorrelated, that is, there is no systematic within-person error. Howe (1985) has conducted simulations of various approaches that use dual responses for all subjects and

has confirmed that the additional information improves the estimate of the relative risk and the statistical power of tests under most circumstances.

CORRECTION FOR MEASUREMENT ERROR IN CONFOUNDING VARIABLES

As demonstrated by Greenland (1980) and Kupper (1984), error in the measurement of confounding variables can distort relative risk estimates in any direction. Kupper has provided a method to adjust partial correlation coefficients for such errors; the details are beyond the scope of this book. Analogous methods for relative risk adjustments are under development but are not yet readily available.

SUMMARY

A variety of methods exist to correct epidemiologic measures of association for error in the measurement of exposure. Each method requires assumptions that are rarely perfectly satisfied. For example, many methods require that errors be strictly due to random within-person variation. Methods based on a validation substudy assume that the true measure is indeed true. These assumptions are rarely perfectly correct, and it is typically difficult to evaluate whether deviations from these assumptions are serious from a practical standpoint. Despite these limitations, careful use of these adjustment procedures should provide better estimates of the quantitative relationships between nutritional factors and disease than analyses that ignore the effects of measurement error altogether.

REFERENCES

Anderson, S. A. (1986). *Guidelines for the Use of Dietary Intake Data.* Federation of American Societies for Experimental Biology, Bethesda, Md.: Life Sciences Research Office.

Armstrong, B. (1985). Measurement error in the generalized linear model. *Commun. Statist. Simula Computa 14,* 529–544.

Barron, B. A. (1977). The effects of misclassification on the estimation of relative risk. *Biometrics 3,* 414–418.

Beaton, G. H., J. Milner, P. Corey, et al. (1979). Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *Am. J. Clin. Nutr. 32,* 2546–2259.

Byars, D. and M. Gail (in press). Proceedings of a workshop on measurement error. *Statistics in Medicine.*

Carroll, R. J., C. H. Spiegelman, K. K. Gordon, Lan, K. T. Bailey, and R. D. Abbott (1984). On errors-in-variables for binary regression models. *Biometrika 71,* 19–25.

Donner, A. and J. J. Koval (1980). The estimation of intraclass correlation in the analysis of family data. *Biometrics 36,* 19–25.

Greenland, S. (1980). The effect of misclassification in the presence of covariates. *Am. J. Epidemiol. 112*, 564–569.

Howe, G. R. (1985). The use of polytomous dual response data to increase power in case control studies: An application to the association between dietary fat and breast cancer. *J. Chron. Dis. 38*, 663–670.

Kaldor, J. and D. Clayton (1985). Latent class analysis in chronic disease epidemiology. *Statistics in Medicine 4*, 327–335.

Kleinbaum, D. G., L. L. Kupper, and H. Morganstern. (1982). *Epidemiologic Research.* Belmont, Calif.: Lifetime Learning Publication.

Kupper, L. (1984). Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. *Am. J. Epidemiol. 120*, 643–648.

Liu, K., J. Stampler, A. Dyer, et al. (1978). Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J. Chron. Dis. 31*, 399–418.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Am. Stat. Assoc. 54*, 173–205.

Marshall, J. R. and S. Graham (1984). Use of dual response to increase validity of case-control studies. *J. Chron. Dis. 37*, 125–136.

Mattson, E. H., B. A. Erickson, and A. M. Kligman (1972). Effect of dietary cholesterol on serum cholesterol in man. *Am. J. Clin. Nutr. 25*, 589–94.

Miettinen, 0. (1976). Estimability and estimation in case-referent studies. *Am. J. Epidemiol. 103*, 226–235.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika 69*, 331–342.

Riggs, D. S., J. A. Guarnieri, and S. Addelman (1978). Fitting straight lines when both variables are subject to error. *Life Sciences 22*, 1305–1360.

Rosner, B. and W. C. Willett (1988). Interval estimates for correlation coefficients corrected for within-person variation: Implications for study design and hypothesis testing. *Am. J. Epidemiol. 127*, 377–386.

Rosner, B. A., W. C. Willett, and D. Speigelman (1989). Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Statistics in Medicine* (in press).

Snedecor, W. G. (1968). Error in measurement in statistics. *Technometrics 10*, 637–666.

Stefanski, L. A. and R. J. Carroll (1985). Covariate measurement error in logistic regression. *Ann. Statist. 13*, 1335–1351.

Tukey, J. W. (1951). Components in regression. *Biometrics 7*, 33–69.

Walker, A. M. and M. Blettner (1985). Comparing imperfect measures of exposure. *Am. J. Epidemiol. 121*, 783–790.

Walter, S. D. (1984). Commentary on "Use of dual responses to increase validity of case-control studies." *J. Chron. Dis. 37*, 137–142.

Willett, W. C., L. Sampson, M. J. Stampfer, et al. (1985). Reproducibility and validity of a semi-quantitative food frequency questionnaire. *Am. J. Epidemiol. 122*, 51–65.

Willett, W. C., R. D. Reynolds, and S. Cottrell-Hoehner, et al. (1987a). Validation of a semi-quantitative food frequency questionnaire: Comparison with a one-year diet record. *J. Am. Diet. Assoc. 87*, 43–7.

Willett, W. C., M. J. Stampfer, G. A. Colditz, et al. (1987b). Dietary fat and risk of breast cancer. *N. Engl. J. Med. 316*, 22–28.

In the next three chapters, concepts described in the earlier sections are discussed in relation to specific topics of current epidemiologic interest: vitamin A intake and lung cancer, dietary fat and breast cancer, and the diet–heart hypothesis. Since space precluded a detailed review of all potentially important relationships between dietary factors and specific diseases, these topics were selected largely because a sufficient literature exists for each to illustrate some of the important principles in nutritional epidemiology. The substance of these chapters may become outdated rather quickly as these are areas of active investigation. Nevertheless, the lessons learned trying to unravel these complex relationships should be useful to investigators and consumers of the scientific literature who address these and other problems in the future.

13

Vitamin A and Lung Cancer

It has long been recognized that vitamin A plays a central physiologic role in the regulation of cell differentiation (Wolbach and Howe, 1925; DeLuca et al., 1972). Because loss of differentiation is a basic feature of malignancy, vitamin A may be related to cancer incidence. In numerous animal studies, naturally occurring preformed vitamin A and synthetic analogues have inhibited the occurrence of induced tumors and even reversed metaplastic changes (Hill and Grubbs, 1982; Sporn and Roberts, 1983). This inhibitory effect is present even when retinol is administered after the cancer has been induced (McCormick et al., 1982), a feature of major potential epidemiologic and public health importance. Under some laboratory conditions, however, these same compounds can increase the incidence of tumors (Schroder and Black, 1980).

Interest in vitamin A as a potential inhibitor of human cancer increased substantially with a report by Bjelke (1975) based on the follow-up of 8278 Norwegian men who had earlier completed a mailed dietary questionnaire. After adjusting for the effects of cigarette smoking, Bjelke observed that the rate of lung cancer among men whose calculated intake of vitamin A was above average was only one-third that of men with intakes below average (Table 13–1).

The interpretation of Bjelke's important finding was complicated by the diversity of vitamin A sources. Natural preformed vitamin A, frequently referred to as retinol even though it is usually consumed in the form of retinyl esters, is found only in foods from animal sources. Plants, principally green leafy

Table 13–1. Age-adjusted rates of lung cancer among men with high versus low vitamin A index

	Rate of lung cancer[a]		Relative risk
	Vitamin A index		Vitamin A index
	≤5	≥5	≥5 vs <5
Cigarette smoking status			
Ever smoked	10.6	4.2	0.40[b]
Current smoker, >20 cigarettes/day	21.0	7.4	0.35
Current smoker, 1–19 cigarettes/day	12.8	5.7	0.44
Ex-smoker	6.1	1.5	0.25
Never smoked	1.1	1.2	1.01
Total, smoking adjusted	7.3	2.8	0.38[c]

[a]Age-adjusted 5-year cumulative incidence/1000 men.
[b]$p < 0.05$.
[c]$p < 0.01$.
Data are based on 53 cases occurring during a 5-year follow-up of 8278 Norwegian men.
From Bjelke, 1975.

vegetables and yellow fruits and vegetables, contain not preformed vitamin A but a series of carotenoid compounds, some of which can be metabolized to form retinol, the physiologically active form of vitamin A. Beta-carotene, the most plentiful carotenoid with potential vitamin A activity, is a dimer that can be cleaved after absorption to form two molecules of retinol (Fig. 13–1). The majority of carotenoids in our food supply, such as lycopene that is found in egg yolks, contribute yellow and red colors, but not vitamin A activity to our diet. Some animal products contain modest amounts of beta-carotene and other carotenoids; without them, butter and chicken fat would appear white. Vitamin

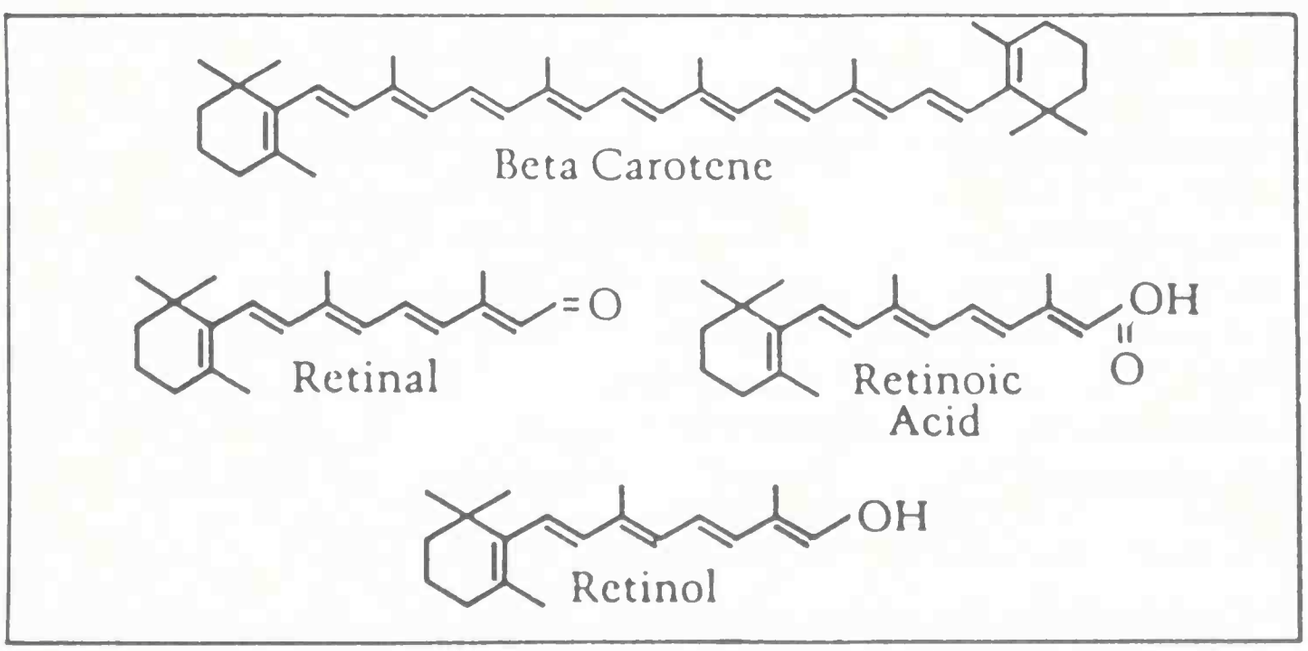


Figure 13–1. Beta-carotene and retinol. Some of the beta-carotene may be cleaved after absorption to form two molecules of retinol, and some circulates unchanged. (Reproduced with permission from Hoffmann-LaRoche.)

Table 13–2. Preformed vitamin A and vitamin A activity from carotenoids in several commonly consumed foods

	Typical serving	Preformed vitamin A	Carotenoids
Major sources of preformed vitamin A			
Liver	3½ oz	52,033	298
Breakfast cereal, typical	½ cup	1,250	0
Margarine	1 pat (5 g)	97	58
Hard cheese	1 oz	183	117
Eggs	one	260	0
Low-fat milk	8 oz glass	499	0
Butter	1 pat (5 g)	95	58
Skim milk	8 oz glass	499	0
Whole milk	8 oz glass	190	117
Major sources of carotenoids			
Carrots	1 carrot	0	19,152
Spinach	½ cup, cooked	0	7,371
Tomatoes	1 fruit	0	1,394
Yellow squash	½ cup	0	3,628
Broccoli	½ cup	0	1,099
Orange juice	6 oz glass	0	310
Mixed peas and carrots	½ cup	0	3,892
Canteloupe	4 oz slice	0	4,304
Tomato paste, sauce	4 oz	0	1,528

Data from total vitamin A activity based on U.S.D.A. Handbook No. 8 (U.S. Department of Agriculture) and provisional tables. For foods containing both preformed vitamin A and carotenoid, the total vitamin A activity has been partitioned based on data from Paul and Southgate, 1978.

A supplements and fortified foods, such as breakfast cereals, are almost always based on preformed vitamin A; only very recently has it been possible for consumers to purchase directly beta-carotene supplements. Foods that provide large amounts of vitamin A activity from preformed vitamin A and carotenoids in the U.S. diet are described in Table 13–2.

Although some carotenoids share common physiologic functions with preformed vitamin A by virture of their potential conversion to retinol, carotenoids may have other actions that are unique. The carotenoids efficiently quench singlet oxygen and free radicals that could otherwise initiate reactions such as lipid peroxidation (Peto et al., 1981; Krinsky and Deneke, 1982). Beta-carotene is a potent free radical scavenger under physiologic conditions of oxygen tension, a characteristic not shared by retinol or other carotenoids (Burton and Ingold, 1984). It is thus important to distinguish between the effects of preformed vitamin A, derived from animal products and supplements, and carotenoids, obtained largely from plant products (Peto, 1983).

ADDITIONAL PROSPECTIVE STUDIES

The original findings of Bjelke were confirmed by a second report with extended follow-up of this cohort (Kvale et al., 1983) that also provided more information on specific foods. The protective effect against lung cancer was primarily attrib-

utable to carrots and other vegetables with some additional contribution from milk. This study thus provided stronger evidence for a beneficial effect of carotenoid sources of vitamin A than for preformed vitamin A.

The first epidemiologic study in which the independent effects of carotenoids and preformed vitamin A were formally examined was the 19-year follow-up of 2107 men enrolled in the Western Electric Study (Shekelle et al., 1981). Preformed vitamin A intake was not related to the incidence of lung cancer, which was diagnosed among 33 men. In striking contrast, intake of vitamin A from carotenoids was strongly associated with lower risk of this disease (Fig. 13–2). No significant relationship was observed between intake of either form of vitamin A and cancers other than lung in this cohort, but the number of cases with these cancers was small.

The relationships of vitamin A intake with risk of lung cancer have been examined in two other prospective studies, both of which had very limited dietary information. Hirayama (1979) examined the association between intake of green and yellow vegetables, assessed with a single question, and the occurrence of 807 lung cancer deaths during 10 years of follow-up among a cohort of 265,118 Japanese men and women. A reduced risk of lung cancer was observed for both smokers and nonsmokers who ate these vegetables daily compared with those who did not. The relative risk ranged from 0.4 to 0.8, depending on gender and smoking habits.

The other prospective study of vitamin A and lung cancer was based on 1

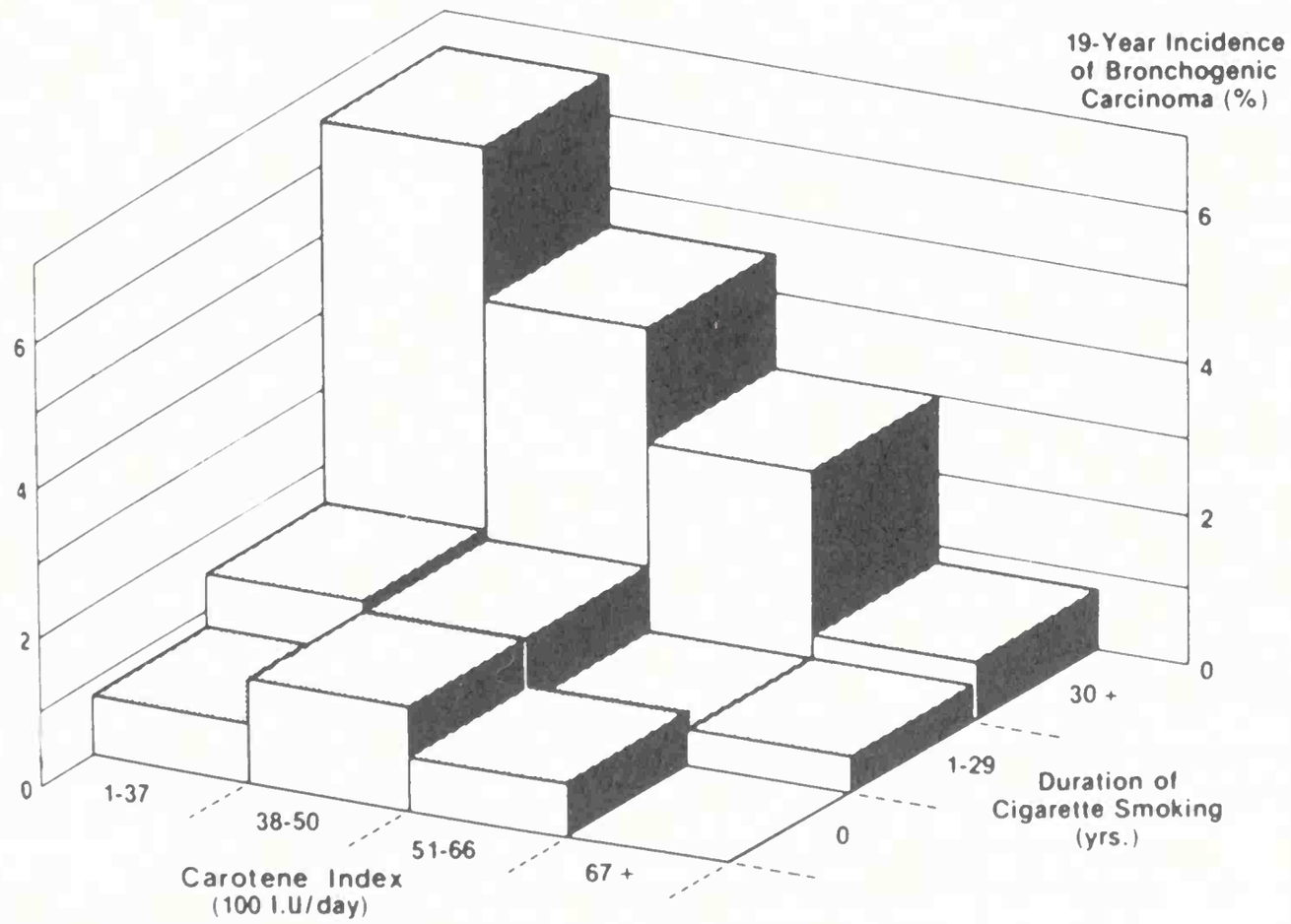


Figure 13–2. Association of carotene index and cigarette smoking with 19-year incidence of bronchogenic carcinoma. (From Shekelle et al., 1981, reproduced with permission.)

million men and women enrolled in the American Cancer Society cohort; 2952 deaths due to lung cancer occurred during 10 years of follow-up (Long-de and Hammond, 1985). The details of analysis were quite incomplete; however, a strong protective effect was observed for those who consumed green salad and fruit or fruit juice five to seven times per week compared with those who consumed these foods fewer than three times per week. Although the dietary information in both this and the Japanese investigation was extremely crude, the prospective nature of the studies reduces the possibility of methodologic bias. Although both studies provide support for the hypothesis that carotenoids reduce the risk of lung cancer, the very limited nature of the data does not provide opportunity to explore alternate hypotheses, for example, that vitamin C may be the responsible factor. Furthermore, they do not address the potential effect of preformed vitamin A.

CASE-CONTROL STUDIES

The relationship between vitamin A intake and risk of lung cancer has been examined in more than seven case-control studies, which are summarized in Table 13–3. An explicit distinction between preformed vitamin A and carot-

Table 13–3. Case-control studies of vitamin A intake and lung cancer

Study	Populations	Number of cases	Relative risks for high vs low intake categories		
			Total vitamin A	Preformed vitamin A	Carotenoid vitamin A
MacLennan et al. (1977)	Chinese men and women in Singapore	233	—	—	0.6–0.7
Mettlin et al. (1979)	New York men	292	0.4	—	—
Gregor et al. (1980)	English men and women	100	0.5 (men) 1.9 (women)	— —	— —
Hinds et al. (1984)	Hawaiian men and women	364	0.7	—	—
Ziegler et al. (1984)	New Jersey white men	763	1.1	1.3	0.8
Samet et al. (1985)	New Mexico men and women	322 Anglos 125 Hispanics	0.6 1.1	1.1 1.7	0.7 1.1
Wu et al. (1985)	U.S. women	220	— —	0.8[a] 1.0[b]	0.4[a] 0.7[b]
Byers et al. (1987)	U.S. men and women	296 men 154 women	0.7 0.8	— —	0.6 0.8
Pastorino et al. (1987)	Italian women	47	—	0.4	0.3

[a]Adenocarcinoma.

[b]Squamous cell carcinoma.

enoid intake was made in the studies of Ziegler and co-workers (1984), Samet and co-workers (1985), Wu and co-workers (1985), and Pastorino and co-workers (1987). In these investigations a statistically significant protective effect was observed only for carotenoid sources of vitamin A; in the study of Samet and colleagues this was limited to Anglo and not Hispanic subjects. In the investigation of MacLennan and colleagues (1977) the inverse relationship was attributable to green leafy vegetables, which would contain carotenoids rather than preformed vitamin A. Similarly, in a study by Pisani and colleagues (1986) a strong inverse association was found for carrots, although intakes of carotenoids and preformed vitamin A were not calculated. In the report of Mettlin and colleagues (1979) the protective effect of total vitamin A was primarily due to carrots and milk; because dairy products generally contain both carotenoids and preformed vitamin A, this finding would be consistent with protection entirely attributable to carotenoids, but an effect of preformed vitamin A cannot be excluded. Hinds and colleagues (1984) provided data for total vitamin A and carotene; because the protective effect of carotene was slightly greater than for total vitamin A and the large majority of vitamin A in the diet of these subjects was in the form of carotene, it can be concluded that preformed vitamin A had little relationship with risk of lung cancer. Nearly identical findings were observed in a large study by Byers and co-workers (1987). The small English study of Gregor and co-workers (1980) provided weak evidence of a protective effect for preformed vitamin A as the inverse association among men was attributable to intake of liver and vitamin supplements; however, this was nearly balanced by a non-significant positive association among the small number of women. The apparent protective effect of preformed vitamin A in the small Italian study of Pastorino and colleagues (1987) did not attain statistical significance.

As a whole, the case-control investigations are remarkably consistent with the prospective studies in providing support for a protective relationship between carotenoid sources of vitamin A and the occurrence of lung cancer. In contrast, they suggest that intake of preformed vitamin A has little or no relationship with the incidence of lung cancer within the range of intakes that were studied. Associations of carotenoid sources with other cancers have been observed in some case-control studies but consistent findings have not emerged; a review of this literature is outside the scope of this chapter.

STUDIES OF BLOOD VITAMIN A LEVELS AND RISK OF LUNG CANCER

Using a very different approach, a number of investigators have examined the relationship between blood levels of vitamin A and lung cancer. The specific measurements that have been employed are discussed in further detail in Chapter 9; these have included serum or plasma levels of retinol, retinol binding protein, total carotenoid, and beta-carotene.

Because the relationships between biochemical measurements and disease are usually much easier to investigate using a case-control study rather than a

prospective design, it was natural that the early investigations of vitamin A and lung cancer used this approach. For example, Atukorala and colleagues (1979) reported levels of serum total vitamin A and beta-carotene (probably this was actually total carotenoids as they used an older spectrophotometric assay) among 26 patients with newly diagnosed lung cancer and 21 control subjects. They found that vitamin A levels were approximately 25 percent lower in the cases ($p < 0.001$); however, the difference in carotenoid levels was less and did not reach statistical significance. As the authors pointed out, this type of study could not distinguish the possibility that the low vitamin A levels might be a consequence of the tumor rather than a cause of the malignancy.

To avoid the possibility that the cancer might affect the level of vitamin A or other biochemical indicators, the specimens should be collected before the occurrence of disease. Although it would be ideal to analyze all the specimens at the time of collection and then follow the cohort to ascertain the occurrence of cancer, this is usually prohibitive in terms of cost and time. Thus, a number of investigators have conducted nested case-control studies using collections of sera that had been obtained and stored for other purposes, usually as part of studies of cardiovascular disease. Wald and colleagues (1980) followed a cohort of about 16,000 men and identified 86 cancers that occurred within 3 years of having provided a blood specimen that was frozen and stored. Sera for each case of cancer and two control subjects from this population matched by age, smoking habit, and date of blood collection were retrieved and analyzed for retinol. Overall, serum retinol levels in these specimens, collected before diagnosis, were 7 percent lower among the case subjects compared with levels among the controls (Table 13–4). This difference was statistically significant and translated into a dose-response relationship across quintiles with a relative risk of 0.6 for men in the highest category of serum retinol compared with those in the lowest category. At approximately the same time, Kark and colleagues (1981) reported data from a similar study conducted in Evans County, Georgia. They found that the prediagnostic retinol levels in the cases averaged 12 percent lower than those of the control subjects, which corresponded to a dose-response relationship with a relative risk of 0.2 for those in the highest compared with the lowest quintile. In both the Wald and Kark studies the difference between cases and controls was particularly large for lung cancer, which seemed apparently consistent with the case-control data on vitamin A intake. These findings generated great interest in the area of cancer research as it appeared that retinol might be to cancer what serum cholesterol is to the field of cardiovascular epidemiology.

Despite the intitial optimism that serum retinol levels might be a strong and potentially modifiable risk factor for cancer overall and lung cancer in particular, further studies did not confirm the original findings. In a preliminary report from Switzerland based on sera that were actually analyzed at the time of specimen collection, Stahelin and colleagues (1982) did not find lower levels of retinol among men who subsequently developed lung or all cancers combined (120 total cases). Willett and colleagues (1984) used a nested case-control design and sera that had been collected as part of a national multicentered trial of hypertension treatment to evaluate the hypothesis that low serum retinol levels are associated with an increased risk of cancer. No association was seen with the

Table 13–4. Studies of retinol levels in prospectively collected sera in relation to risk of lung cancer

Study	Population	N of Cases	Mean retinol level (μg/dl)			Relative risk, high vs low category
			Cases	Controls		
Wald et al. (1980)	English men	14 lung cancers	56.1	68.7	$p < 0.005$	-
		86 total cancers	64.2	68.7	$p < 0.025$	0.5
Kark et al. (1981)	Evans Co. men and women	12 lung cancers	(cases 9.1 μg/dl lower)			-
		85 total cancers	41.3	46.9	$p = 0.003$	0.2
Willett et al. (1984)	U.S. men and women, hypertensive	17 lung cancers	(cases 7.4 μg/dl higher)			-
		111 total cancers	67.3	68.7		1.1
Stahelin et al. (1984)	Swiss	35 lung cancers	286 iu/dl	280 iu/dl		-
		(other sites displayed separately)				
Peleg et al. (1984)	Evans Co. men and women	17 lung cancers	(cases 6.4 μg/dl lower)			-
		135 total cancers	48.9	48.9		1.2
Salonen et al. (1985)	Finnish men and women	15 lung cancers	48.0	56.8	$p < 0.05$	-
		51 total cancers	48.3	52.4		1.1
Nomura et al. (1985)	Japanese–Hawaiian men	74 lung cancers	63.8[a]	59.6[a]		-
Friedman et al. (1986)	California men and women	151 lung cancers	82.2	82.4		0.8
Menkes et al. (1986)	Washington Co. men and women	99 lung cancers	60.6	61.3		0.9
Wald et al. (1986)	English men	41 lung cancers	65.6	66.3		-
		227 total cancers	67.0	68.8		0.8

[a]Median value.

overall incidence of cancer, and for lung cancer the levels were actually somewhat higher for cases than controls. A similar lack of any protective relationship was seen for retinol-binding protein. In a small study conducted in Finland, Salonen and colleagues (1985) found no relationship between retinol levels and overall risk of cancer; however, the levels among the lung cancer cases were somewhat lower than among the control subjects. More recently, three studies have been conducted with a sufficient number of total cancers to examine lung cancer alone with substantial power (Nomura et al., 1985; Friedman et al., 1986; Menkes et al., 1986). In all three of these studies, the retinol levels of subjects who later developed lung cancer were nearly identical to those who remained free of cancer. In addition, Peleg and colleagues (1984) have published additional follow-up data from the Evans county study and Wald and colleagues (1986) have reported on further experience within the English cohort. In both these extensions of the two original studies in which a protective relationship was found, there no longer existed any material association between serum retinol and incidence of lung cancer. Indeed, during the extended follow-up in the

English study, the relationship between retinol level and risk of lung cancer and total cancers was actually positive, although not significantly so. The body of evidence relating serum retinol levels to risk of lung cancer as well as overall cancer incidence is thus now overwhelmingly null, although the possibility cannot be excluded that an association exists with some less common specific malignancy.

Now that the initial excitement surrounding serum retinol and cancer has subsided, it may be informative to consider whether preventable methodologic weaknesses were responsible for the initial inverse relationships. Wald and colleagues (1986) have considered this issue in detail (Table 13–5). In their second paper based on additional follow-up, they confirmed with new data that an inverse relationship existed during the first 3 years following collection of the blood samples. As noted above, however, the inverse relationship did not persist beyond 3 years and actually reversed its direction. Wald and co-workers concluded that the initial finding was not likely to be due to chance; rather it is more likely that preclinical malignancy, which would have been present at the time of blood collection for tumors diagnosed during the early follow-up period, was reducing the level of retinol by some unknown metabolic mechanism. This is analogous to the apparent inverse relationship between serum cholesterol and cancer risk, which is largely manifested during the early follow-up after blood collection (Rose and Shipley, 1980). Although it is not surprising that case-control studies of serum markers of nutritional status and cancer should provide misleading results, a sobering point raised by these studies of serum retinol and cancer is that even prospective studies may not be prospective enough. No single approach can provide absolute assurance that this problem will not be repeated. Extending the follow-up period, careful screening of subjects at the time of blood collection to eliminate prevalent disease, using early, superficial, or premalignant endpoints, and employing biochemical indicators that are less sensitive to recent metabolic or dietary changes, such as nails or subcutaneous fat, however, will reduce the likelihood of bias due to existing malignancies.

Table 13–5. Relative risks of cancer (all sites) according to quintile of serum retinol in preliminary report and during different follow-up intervals in the extended study

	Relative risks (vs. lowest quintile)			
Quintile of retinol	Preliminary study (86 cases)	Second study <1 yr (66 cases)	Second study 1–2 yr (45 cases)	Second study 3+ yr (116 cases)
1 low	1.00	1.00	1.00	1.00
2	0.77	0.38	1.26	0.86
3	0.73	0.49	0.49	1.35
4	0.48	0.14	0.33	0.82
5 high	0.45	0.32	0.25	1.87
Test for trend in relative risks	$p < 0.025$	$p < 0.01$	$p < 0.025$	NS

NS = not significant.

From Wald et al., 1980 and 1986.

Although preclinical disease may have distorted the relationship between serum retinol and cancer in the initial study of Wald and colleagues, this phenomenon is not likely to explain the findings of Kark and colleagues, because subjects were followed up to 14 years. In this study, however, the sera for cases were handled in a systematically different manner from those of controls as the specimens for cases but not controls had been thawed and refrozen as part of another investigation. The authors were aware of this potential problem and conducted a substudy to learn whether thawing and refreezing specimens would reduce the retinol levels; they detected no such effect. It is impossible, however, to reproduce exactly the procedures that the case sera had experienced, leaving concern that this differential treatment explained the observed inverse association.

The underlying issue in these two early studies is that relationships between nutritional factors and disease are likely to be exquisitely sensitive to methodologic distortion as small differences in mean values, which can be due to seemingly minor systematic biases, can translate into important relative risks. It is reassuring, however, that the nested case-case control studies of serum retinol and cancer that have not been limited to only a short follow-up period and in which case and control specimens were handled identically have been remarkably consistent.

In retrospect, it also appears that the apparent convergence between the studies of vitamin A intake and serum retinol in relation to lung cancer was based on a misunderstanding of the relationship of vitamin A intake and serum retinol (see Chapter 9). Although serum retinol levels are clearly depressed with intakes sufficiently low to produce signs of clinical deficiency, over the range of diets in generally well-nourished populations, serum retinol levels are only very minimally related to vitamin A intake. For example, in one study in which 25,000 IU of preformed vitamin A (more than five times the usual dietary intake) was taken as a daily supplement for 8 weeks, no detectable increase in blood level was observed (Willett et al., 1983). Similarly, large supplements of beta-carotene had no material effect on blood levels of retinol. In another randomized trial among women with initially low values of plasma retinol, a daily supplement of 10,000 IU increased blood levels by only 9 percent, which was marginally statistically significant (Willett et al., 1984). Therefore, the failure to observe an association between serum retinol level and lung cancer incidence reveals little about the relationship between dietary vitamin A and this disease and does not preclude an effect of intake of this vitamin on risk of cancer.

NESTED CASE-CONTROL STUDIES OF BLOOD CAROTENE LEVELS AND LUNG CANCER

Based on the data from studies of dietary intake, it would be expected that blood levels of beta-carotene would be related to risk of lung cancer. This relationship has been more difficult to study than the association with serum retinol for two reasons. First, assays to measure beta-carotene in large numbers of blood specimens have only recently become available; the older spectrophotometric

methods could not distinguish beta-carotene from the other more plentiful carotenoids. Second, beta-carotene is far less stable in frozen sera than retinol so that ultra-low temperatures (e.g., −80°C) are needed to avoid degradation if specimens are to be held for more than a few months (see Chapter 9). Because a number of serum banks did not use ultra-low conditions, many studies were not able to measure this nutrient.

At the time that the nested case-control study reported by Willett and colleagues (1984) was conducted, an assay was not available for beta-carotene, which is only about 10 to 15 percent of total serum carotenoids, so that total carotenoids were measured instead; no significant association was observed either for lung or total cancers. Subsequently, Stahelin and colleagues (1984) reported significantly lower prediagnostic levels of serum beta-carotene in 35 lung cancer patients than in controls. After adjusting for age and the number of cigarettes smoked daily, Nomura and colleagues (1985) observed a significant inverse relation between serum beta-carotene levels and incidence of lung cancer; the relative risk was 0.5 for men in the highest compared with the lowest quintile. In a similar analysis, Menkes and colleagues (1986) found that higher serum beta-carotene levels were associated with a lower risk of lung cancer among the Washington County cohort; the relative risk for subjects in the extreme quintiles was also 0.5.

In studies that have been able to measure beta-carotene in bloods collected before the diagnosis of lung cancer, an inverse relationship has thus been consistently observed. Although this adds support to the hypothesis that higher intakes of beta-carotene reduce the risk of lung cancer, these findings also need to be interpreted cautiously. Studies based on biochemical measurements of beta-carotene have little advantage over those based on dietary intake with respect to control of potentially confounding factors. For example, in the dietary intake studies, it is possible that some factor in the green and yellow fruits and vegetables other than beta-carotene is responsible for the reduced risk of lung cancer. Similarly, it is possible that some other aspect of life-style among persons who consume higher amounts of these fruits and vegetables is responsible for the lower cancer incidence. These same limitations apply to the biochemical studies because the higher blood levels of beta-carotene could be only incidental to the increased intake of other protective factors in these plant products or to protective behavioral characteristics associated with diet. Although it is useful to obtain similar answers using different approaches, the value of these studies lies more in their prospective nature rather than their means of measuring exposure. Indeed, even the simple questionnaire employed in the study of Bjelke provided more opportunity than the biochemical studies to evaluate alternative explanations for the inverse association with "vitamin A" as these authors could calculate an index of vitamin C intake and demonstrate that it did not account for the association with vitamin A. If the biochemical assay for beta-carotene also provides independent values for other carotenoids, these could be similarly examined as potential determinants of disease. Studies using biochemical markers can sometimes be informative when dietary intake studies are not; if beta-carotene were difficult or impossible to measure by questionnaire, then the biochemical marker could provide unique information. Because carotene intake

derived from a questionnaire is correlated with blood levels of beta-carotene or total carotenoids (Willett et al., 1983; Russell-Briefel et al., 1985; Stryker et al., 1987), the biochemical marker does not provide special information.

The interpretation of data on blood levels of beta-carotene and lung cancer is also complicated by recent suggestions that cigarette smoking may have a metabolic effect that reduces blood levels of beta-carotene. Although blood levels of total carotenoids and beta-carotene have been found in a number of studies to be lower among cigarette smokers than among nonsmokers, it is possible that this is simply the result of reduced intake of fruits and vegetables among smokers. Russell-Briefel and colleagues (1985), however, found that these differences in carotenoid levels could not be accounted for by dietary intake as measured by a simple food-frequency questionnaire. This finding was recently confirmed by Stryker and colleagues (1988), who also found that the slope of the regression line relating carotene intake to plasma beta-carotene was significantly reduced among smokers compared with nonsmokers. If this is confirmed, this implies that smoking has a direct metabolic effect on blood beta-carotene and renders those studies of blood beta-carotene and lung cancer risk, viewed on their own, almost uninterpretable. In theory, these studies have controlled for the confounding effects of smoking by adjusting for the reported number of cigarettes smoked daily. In reality, this is a rather crude measure of the actual physiologic impact of cigarette smoking as persons smoking exactly the same number of cigarettes per day can differ in unmeasurable details of the cigarette smoked, depth of inhalation, and the anatomic structure of their lungs. To the extent that these factors determine the concentration of tobacco combustion products in lung tissue and thus influence both the risk of lung cancer and blood levels of beta-carotene, confounding due to the actual physiologic intensity of smoking is not fully controllable in epidemiologic analyses. This problem could be particularly intractable because of the overwhelmingly strong relationship between cigarette smoking and lung cancer.

Fortunately, evidence for a relationship between beta-carotene and lung cancer is not solely based on the biochemical data. The consistent observation between intake of carotenoids and this disease strongly suggests that the higher risk of lung cancer among persons with low serum beta-carotene levels is not merely due to a metabolic effect of smoking on the blood levels, although it is possible that the magnitude of the relationship with serum levels still may be exaggerated. This issue illustrates one of the potential limitations of biochemical markers of diet. It has been frequently said that an advantage of such markers is that they provide a measure of exposure that is more proximal to the disease process than measurements based on dietary intake. If the question being addressed is the relation of dietary intake to risk of disease, however, then the biochemical indicators are less direct measures of exposure than diet itself. As such, they may be influenced by nondietary factors that can add extraneous variation. More seriously, the biochemical indicators are subject to the same sources of confounding as the dietary intake measures, plus additional sources of confounding that affect the biochemical levels independent of dietary intake. Thus these two general approaches of measuring exposure should be viewed as complementary, rather than one being a substitute for the other.

FURTHER CONSIDERATIONS

The remarkable consistency of the association between carotenoid intake and lung cancer seen in the case-control and cohort studies as well as those investigations based on prediagnostic blood specimens strongly suggests that this relationship is not due to chance or methodologic artifact. What remains less certain is whether this finding represents a true causal effect of beta-carotene or a confounding effect of another dietary or nondietary factor associated with carotene intake. It is particularly difficult to exclude the possibility that another component of certain fruits and vegetables, such as a carotenoid other than beta-carotene or an indole compound (Wattenberg and Loub, 1978), is the truly protective factor. The available data are, therefore, probably best interpreted as providing strong support for a protective effect of fruits and vegetables against lung cancer.

Ideally, the issue of confounding is best resolved in a randomized trial, as discussed later. Additional insight, however, can be gained by examining the relationhips between individual food items and risk of lung cancer. The observations that carrots (Mettlin et al., 1979; Kvale et al., 1983) as well as green leafy vegetables (MacLennan et al., 1977; Ziegler et al., 1984) are associated with reduced risk provide some additional evidence that beta-carotene is the active agent as these distinctly different types of vegetables both have unusually high concentrations of this nutrient. Unfortunately, the data from several studies have not been fully exploited by carefully examining the associations of specific foods or food groups with risk of lung cancer. In conducting such analyses it is important to control the effects of one food for use of other foods related to risk of disease as their consumption is typically correlated.

Even though the evidence relating beta-carotene intake to lung cancer risk is not conclusive, the nature of the dose-response relationship in existing data would be of substantial interest. Unfortunately, the available data do not allow a detailed examination of this issue for several reasons. Many of the studies were not designed to examine this relationship so that the questions used to assess carotene intake are quite incomplete. Furthermore, issues of nutrient composition are particularly unsettled in the case of carotenoids due to changing laboratory analytic methods and changes in the content of some of the foods themselves. This is reflected in recently updated values for carrots, in which the vitamin A activity has been increased by a factor of more than two (Humble et al., 1985). In addition most food composition tables contain data on vitamin A activity (in international units) rather than absolute amounts of specific carotenoids. Various assumptions have been used regarding the efficiency of conversion of these carotenoids to retinol. The efficiency of conversion, however, is likely to be variable among different persons, in part depending on existing vitamin A status. Moreover, none of the questionnaires used in the investigations reported to date have been subjected to validation studies to quantify their measurement error. It is, therefore, extremely difficult to specify any quantitative relationship between beta-carotene intake and risk of lung cancer.

Even if the quantitative relationship between intake and risk cannot be specified and we are left with a measure of relative intake, the shape of the dose-response curve would be of substantial interest. If only individuals with lowest intake are at elevated risk, it would not be expected that those with average intake would benefit from increasing dietary intake or consuming supplements. If the dose-response relationship was essentially linear, then most persons might potentially benefit from higher intake. At present, the published studies generally appear to be most consistent with a linear dose-response relationship; however, this issue clearly deserves more detailed examination.

In principle, the studies based on biochemical analyses of blood levels of beta-carotene may provide a more standardized measure of exposure and thus better information about dose-response relationships. Only one study (Nomura et al., 1985) provided quantitative information on the relation between intake and risk of disease; for beta-carotene levels greater than 57 μg/dl the relative risk was 0.45 (95% confidence interval = 0.16 − 1.25) compared with levels less than 15 μg/dl. Two reports provided information by quintiles of serum beta-carotene level; for increasing quintiles the relative risks in the study of Nomura and colleagues were 1.0, 0.67, 0.83, 0.42, 0.45 and in the study of Menkes and colleagues (1986) were 1.0, 0.83, 0.56, 0.57, and 0.45. Both of these studies are consistent with a linear dose response relationship. Because of the number of lung cancer cases, however, the confidence intervals for each quintile include 1.0; indeed, the overall trend in both studies was barely statistically significant ($p = 0.04$ in both). Because these estimates of relative risk are quite unstable, the quantitative aspects and even the shape of the dose-response relationship remain to be defined with precision.

Some evidence has suggested that the relationship between carotenoid intake and lung cancer may vary according to the histologic type of the disease and the gender of the subject. In three initial studies, an inverse association was observed only for squamous cell cancer or small cell carcinoma (Kvale et al., 1983; Byers et al., 1984; Zeigler et al., 1984); however, the power to detect an association with adenocarcinoma was not high. In addition, it was suggested by Hinds and colleagues (1984) and Gregor and colleagues (1980) that a protective association may apply to men but not women. In a recent study among women, however, that included 149 cases of adenocarcinoma of the lung, Wu and colleagues (1985) observed a strong protective relationship for this histologic type. The weight of evidence suggests that the protective effect applies to both men and women (MacLennan et al., 1977; Hirayama, 1979; Samet et al., 1985).

RANDOMIZED TRIALS

The most rigorous test of the hypothesis that beta-carotene or another specific form of vitamin A reduces lung cancer risk would be a randomized trial. Compared with trials of other dietary factors this is relatively easy to implement as the active agent can be formulated as a capsule that can be compared with a placebo in a double-blind manner. This is somewhat less straightforward than a

typical drug trial, however, as all participants are consuming some of the active agent and have ready access to more through their diet and, recently, as a supplement from their local health food store. Because the effect of a supplement may well be greatest among persons with low initial dietary intakes of carotene, it would be important to at least measure dietary intake at baseline and, perhaps, to screen potential participants and enroll only those with lower intake. Further considerations for design of such trials are discussed elsewhere (Stampfer et al., 1988).

As of this time no trials of beta-carotene or other forms of vitamin A for the prevention of lung cancer have been reported. Several trials have been started, however, including a large study among men at high risk of lung cancer in Finland (Albanes et al., 1986). In addition, a large study of beta-carotene for the prevention of cancer at all sites combined is being conducted among approximately 22,000 U.S. physicians (Hennekens and Eberlein, 1985). Because physicians are at relatively low risk of lung cancer due to their reduced prevalence of cigarette smoking, it is not clear whether a sufficient number of lung cancers will occur to examine the effect of beta-carotene for this site in particular.

Randomized trials of beta-carotene in relation to the incidence of other cancers or premalignant conditions are also relevant to the hypothesis regarding lung cancer. If an anti-cancer effect is observed for other sites, the likelihood would be increased that the consistent inverse association in the observational studies is causal. Stich and colleagues (1984) have reported that beta-carotene reduces the occurrence of micronuclei, thought to be a premalignant change, in the buccal mucosa of betel nut chewers. More than a dozen other randomized trials of beta-carotene in relation to the occurrence of other malignant or premalignant conditions are now under way; thus, considerably more information will become available over the next several years on the anti-cancer effect of this nutrient.

Although many experiments have been conducted addressing the effect of preformed vitamin A and its synthetic analogues on the occurrence of tumors in animals, beta-carotene has been studied far less extensively in this manner. A clear demonstration of anti-cancer activity would also enhance the causal interpretation of the epidemiologic observations. Beta-carotene, as well as other carotenoids without potential vitamin A activity, protect mice against skin tumors induced by ultraviolet light and chemical carcinogens (Matthews-Roth, 1982). It is not clear, however, whether this represents a specific protective effective against ultraviolet light, which beta-carotene is known to possess (Matthews-Roth et al., 1977). Beta-carotene reduced colon tumors in one experiment (Temple and Basu, 1987) but not in another (Colacchio and Memoli, 1986), and inhibited oral cancers in hamsters when applied topically (Swartz and Sklar, 1987).

SUMMARY

The inverse relationship between intake of vegetables and fruits and risk of lung cancer, which has been found in case-control and cohort studies using both ques-

tionnaire and biochemical measurements of intake, represents one of the best established associations in the field of nutritional epidemiology. Whether this is a causal effect of beta-carotene remains considerably less clear. In principle, dietary advice could be given while awaiting identification of the active agent in these fruits and vegetables; indeed the National Research Council has recommended an increased intake of fruits and vegetables (Committee on Diet, Nutrition and Cancer, 1982). It is unclear, however, which fruits and vegetables are beneficial and what should be the optimal intake of these products. The customary advice to eat a variety of fruits and vegetables is reasonable but leaves much to chance. For example, if beta-carotene is actually an active anti-cancer agent, it is quite possible to eat a variety of these products, such as cucumbers, iceberg lettuce, eggplant, and onions, and still have a low intake of beta-carotene.

The ongoing trials may provide definitive information regarding the effect of beta-carotene if a reduction in cancer is seen. If they are null, however, it will be difficult to exclude the possibility that the treatment was not continued for an adequate time or that study populations did not include a sufficient number of persons who were susceptible by virtue of low dietary carotenoid intake to observe an effect of supplementation. We would then be left with our present knowledge that intake of fruits and vegetable is associated with lower risk of lung cancer. It, therefore, remains important to refine our information based on observational data to the greatest degree possible. This can be done in studies using comprehensive dietary questionnaires so that a wide variety of alternate explanations can be considered. Such questionnaires should be subjected to a validation process to evaluate the degree of error in the assessment of diet. Maximal information would be extracted by conducting data analyses for individual foods as well as nutrients. In investigations using biochemical measurements, it will be helpful to measure a wide variety of nutritional factors, including other carotenoids. The most powerful information would be derived from prospective studies that employed both questionnaire and biochemical assessments.

REFERENCES

Albanes, D., J. Virtamo, M. Rautalahti, J. Pikkarainen, P. R. Taylor, P. Greenwald, and O. P. Heinonen (1986). Pilot study: The U.S.–Finland lung cancer prevention trial. *J. Nutr. Growth Cancer 3,* 207–214.

Atukorala, S., T. K. Basu, J. W. Dickerson, D. Donaldson, and A. Sakula (1979). Vitamin A, zinc, and lung cancer. *Br. J. Cancer 40,* 927–931.

Bjelke, E. (1975). Dietary vitamin A and human lung cancer. *Int. J. Cancer 15,* 561–565.

Burton, G. W. and K. U. Ingold (1984). Beta-carotene: an unusual type of lipid antioxidant. *Science 224,* 569–573.

Byers, T., J. Vena, C. Mettlin, M. Swanson, S. Graham (1984). Dietary vitamin A and lung cancer risk. *Am. J. Epidemiol. 120,* 769–776.

Byers, T. E., C. Vena, B. P. Haughey, J. R. Marshall, M. K. Swanson (1987). Diet and lung cancer: findings from the Western New York Study. *Am. J. Epidemiol. 125,* 351–363.

Colacchio, T. A. and V. A. Memoli (1986). Chemoprevention of colorectal neoplasms. Ascorbic acid and beta-carotene. *Arch. Surg. 121,* 1421–1424.

Committee on Diet, Nutrition and Cancer; National Research Council. (1982). *Diet, Nutrition, and Cancer.* Washington, D. C.: National Academy Press.

De Luca, L., N. Maestri, F. Bonanni, and D. Nelson (1972). Maintenance of epithelial cell differentiation: The mode of action of vitamin A. *Cancer 30,* 1326–1331.

Friedman, G. D., W. S. Blaner, J. H. Vogelman, et al. (1986). Serum retinol and retinol-binding protein levels do not predict subsequent lung cancer. *Am. J. Epidemiol. 123,* 781–789.

Gregor, A., P. N. Lee, F.J.C. Roe, M. J. Wilson, and A. Melton (1980). Comparison of dietary histories in lung cancer cases and controls with special reference to vitamin A. *Nutr. Cancer 2,* 93–97.

Hennekens, C. H. and K. Eberlein (1985). A randomized trial of aspirin and beta-carotene among U.S. physicians. *Prev. Med. 14,* 165–168.

Hill, D. L. and C. J. Grubbs (1982). Retinoid as chemopreventive and anticancer agents intact animals. *Anticancer Res. 2,* 111–124.

Hinds, M. W., L. N. Kolonel, J. H. Hankin, et al. (1984). Dietary vitamin A, carotene, vitamin C, and risk of lung cancer in Hawaii. *Am. J. Epidemiol. 119,* 227–237.

Hirayama, T. (1979). Diet and cancer. *Nutr. Cancer 1,* 67–81.

Humble, C. G., J. M. Samet, and B. E. Skipper (1985). The impact of revisions in vitamin A content data on estimates of nutrient content. *Nutr. Research 5,* 175–179.

Kark, J. D., A. H. Smith, and C. G. Hames (1981). Serum vitamin A (retinol) and cancer incidence in Evans County, Georgia. *J.N.C.I. 66,* 7–16.

Krinsky, N. I. and S. M. Deneke (1982). The interaction of oxygen and oxy-radicals with carotenoids. *J.N.C.I. 69,* 205–210.

Kvale, G., E. Bjelke, and J. J. Gart (1983). Dietary habits and lung cancer risk. *Int. J. Cancer 31,* 397–405.

Long-de,W. and E. C. Hammond (1985). Lung cancer, fruit, green salad and vitamin pills. *Chin. Med. J. 3,* 206–210.

MacLennan, R., J. Da Costa, N. E. Day, et al. (1977). Risk factors for lung cancer in Singapore Chinese. *Int. J. Cancer 20,* 854–860.

Matthews-Roth, M. M., M. A. Pathak, T. B. Fitzpatrick, et al. (1977). Beta-carotene therapy for erythropoietic protoporphyria and other photosensitivity diseases. *Arch. Dermatol. 113,* 1229–1232.

Matthews-Roth, M. M. (1982). Antitumor activity of B-carotene, canthaxanthin and phytoene. *Oncology 39,* 33–37.

McCormick, D. L., F. J. Burns, and R. E. Albert (1981). Inhibition of Benz(a)pyrene-induced mammary carcinogenesis by retinyl acetate. *J.N.C.I. 66,* 559–564.

Menkes, M., G. W. Comstock, J. P. Vuilleumier, et al. (1986). Serum beta-carotene, vitamins A and E, selenium, and the risk of lung cancer. *N. Engl. J. Med. 315,* 1250–1254.

Mettlin, C., S. Graham, and M. Swanson (1979). Vitamin A and lung cancer. *J.N.C.I. 62,* 1435–1438.

Nomura, A. M., G. N. Stammermann, L. K. Heilbrun, et al. (1985). Serum vitamin levels and the risk of cancer of specific sites in men of Japanese ancestry in Hawaii. *Cancer Res. 45,* 2369–2372.

Pastorino, U., P. Pisani, F. Berrino, et al. (1987). Vitamin A and female lung cancer: a case-control study on plasma and diet. *Nutr. Canc. 10,* 171–179.

Paul, A. A. and D.A.T. Southgate (1978). McCance and Widdowson's *The Composition of Foods,* 4th edition, M.R.C. Special Report No. 297, Her Majesty's Stationery Office, London.

Peleg, I., S. Heyden, M. Knowles, et al. (1984). Serum retinol and risk of subsequent cancer: Extension of the Evans County, Georgia, study. *J.N.C.I. 73,* 1455–1458.

Peto, R., R. Doll, J. D. Buckley, et al. (1981). Can dietary beta-carotene materially reduce human cancer rates? *Nature 290,* 201–208.

Peto, R. (1983). The marked differences between carotenoids and retinoids: Methodological implications for biochemical epidemiology. *Cancer Surv. 2,* 327–340.

Pisani, P., F. Berrino, M. Macaluso, U. Pastorino, and P. Crosignani (1986). Carrots, green vegetables and lung cancer: A case-control study. *Int. J. Epidemiol. 15,* 463–468.

Rose, G., and M. J. Shipley (1980). Plasma lipids and mortality: A source of error. *Lancet 1,* 523–526.

Russell-Briefel, R., M. W. Bates, and L. H. Kuller (1985). The relationship of plasma carotenoids to health and biochemical factors in middle-aged men. *Am. J. Epidemiol. 122,* 741–749.

Salonen, J. T., R. Salonen, R. Lappetelainen, et al. (1985). Risk of cancer in relation to serum concentrations of selenium and vitamin A and E: Matched case-control analysis of prospective data. *Br. Med. J. [Clin Res] 290,* 417–420.

Samet, J. M., B. J. Shipper, C. G. Humble, et al. (1985). Lung cancer risk and vitamin A consumption in New Mexico. *Am. Rev. Respir. Dis. 131,* 198–202.

Schroder, E. W. and P. H. Black (1980). Retinoids: Tumor preventers or tumor enhancers. *J.N.C.I. 65,* 671–674.

Shekelle, R. B., M. Lepper, S. Liu, et al. (1981). Dietary vitamin A and risk of cancer in the Western Electric Study. *Lancet 2,* 1186–1190.

Sporn, M. B. and A. B. Roberts (1983). Role of retinoids in differentiation and carcinogenesis. *Cancer Res. 43,* 3034–3040.

Stahelin, H. B., E. Buess, F. Rosel, et al. (1982). Vitamin A, cardiovascular risk factors, and mortality. *Lancet 1,* 394–395.

Stahelin, H. B., F. Rosel, E. Buess, et al. (1984). Cancer, vitamins, and plasma lipids: Prospective Basel study. *J.N.C.I. 73,* 1463–1468.

Stampfer, M. J., W. C. Willett, and C. H. Hennekens (1988). Selection of Population. In Moon, T., Micozzi, M., eds.: *Nutrition in Cancer Prevention.* New York: Plenum Publishing.

Stich, H. F., W. Stich, M. P. Rosin, and M. O. Vallejera (1984). The use of the micronucleus test to monitor the effect of vitamin A, beta-carotene, and canthaxanthin on the buccal mucosa of betel nut/tobacco chewers. *Int. J. Cancer 34,* 745–750.

Stryker, W. S., L. A. Kaplan, E. A. Stein, M. J. Stampfer, A. Sober, and W. C. Willett (1988). The relation of diet, cigarette smoking, and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. *Am. J. Epidemiol. 127,* 283–296.

Schwartz, J. and G. Shklar (1987). Regression of experimental hamster cancer by beta-carotene and algae extracts. *J. Oral. Maxillofac. Surg. 45,* 510–515.

Temple, N. J. and T. K. Basu (1987). Protective effect of β-carotene against colon cancer tumors in mice. *J.N.C.I. 78,* 1211–1214.

United States Department of Agriculture, Agriculture Research Service Handbook No. 8 series. Composition of foods: Raw, processed, prepared. Washington, D.C., U.S. G.P.O., 1976.

Wald, N., M. Idle, J. Boreham, and A. Bailey (1980). Low serum-vitamin A and subsequent risk of cancer. Preliminary results of a prospective study. *Lancet 2,* 813–815.

Wald, N., J. Boreham, and A. Bailey (1986). Serum retinol and subsequent risk of cancer. *Br. J. Cancer 54,* 957–961.

Wattenberg, L. W. and W. D. Loub (1978). Inhibition of polycyclic aromatic hydrocarbon-induced neoplasia by naturally occurring indoles. *Cancer Res. 38,* 1410–1413.

Willett, W. C., M. J. Stampfer, B. A. Underwood, F. E. Speizer, B. Rosner, and C. H.

Hennekens (1983). Validation of a dietary questionnaire with plasma carotenoid and alpha-tocoperhol levels. *Am. J. Clin. Nutr. 38,* 631–639.

Willett, W. C., M. J. Stampfer, B. A. Underwood, L. A. Sampson, C. H. Hennekens, J. C. Wallingford, L. Cooper, C. C. Hsieh, and F. E. Speizer (1984). Vitamin A supplementation and plasma retinol levels: A randomized trial among women. *J.N.C.I. 73,* 1445–1448.

Wolbach, S. B. and P. R. Howe (1925). Tissue changes following deprivation of fat-soluble A vitamin. *J. Exp. Med. 42,* 753–777.

Wu, A. H., B. E. Henderson, M. C. Pike, et al. (1985). Smoking and other risk factors for lung cancer in women. *J.N.C.I. 74,* 747–751.

Ziegler, R. G., T. J. Mason, A. Stemhager, et al. (1984). Dietary carotene and vitamin A and risk of lung cancer among white men in New Jersey. *J.N.C.I. 73,* 1429–1435.

14

Dietary Fat and Breast Cancer

The relationship of dietary fat intake with incidence of breast cancer illustrates issues involved in the synthesis of information from animal studies, correlational data based on population groups, studies of individual subjects, and the interpretation of null findings from epidemiologic data. The subject is also of major public health importance because a putative association has been an important rationale for recent recommendations to reduce fat consumption.

Nearly 50 years ago, Tannenbaum showed that the amount of dietary fat could markedly influence the occurrence of mammary tumors in rodents (Tannenbaum and Silverstone, 1953). For decades this knowledge was largely limited to laboratory scientists, and sometimes regarded as a nuisance variable to those investigating more interesting carcinogens. The publication by Armstrong and Doll (1975) of striking correlations among countries between national per capita fat consumption and both incidence and mortality rates of breast cancer attracted widespread interest in the larger scientific community (see Fig. 14–1 for a display of similar data by Carroll, 1975). The international differences in breast cancer rates are particularly great for postmenopausal women, which has been used to support the hypothesis that diet should be most strongly associated with breast cancer among these women (de Waard et al., 1964). The strong enthusiasm for this hypothesis was reflected in the abstract of a case-control study in which a positive, but not statistically significant, relationship was observed between fat intake and risk of breast cancer (Miller et al., 1978):

> The Study has produced evidence of an association between an increased intake of nutrients, especially total fat, in both pre-menopausal and post-menopausal women with breast cancer. Reasons why a weak association might have been anticipated are discussed, and it is concluded that in reality the association is stronger. Furthermore, its consistency with other evidence, both experimental and international, suggests that it is causal.

After reviewing the published data, a committee of the National Research Council issued a provisional recommendation that the fat content of the U.S.

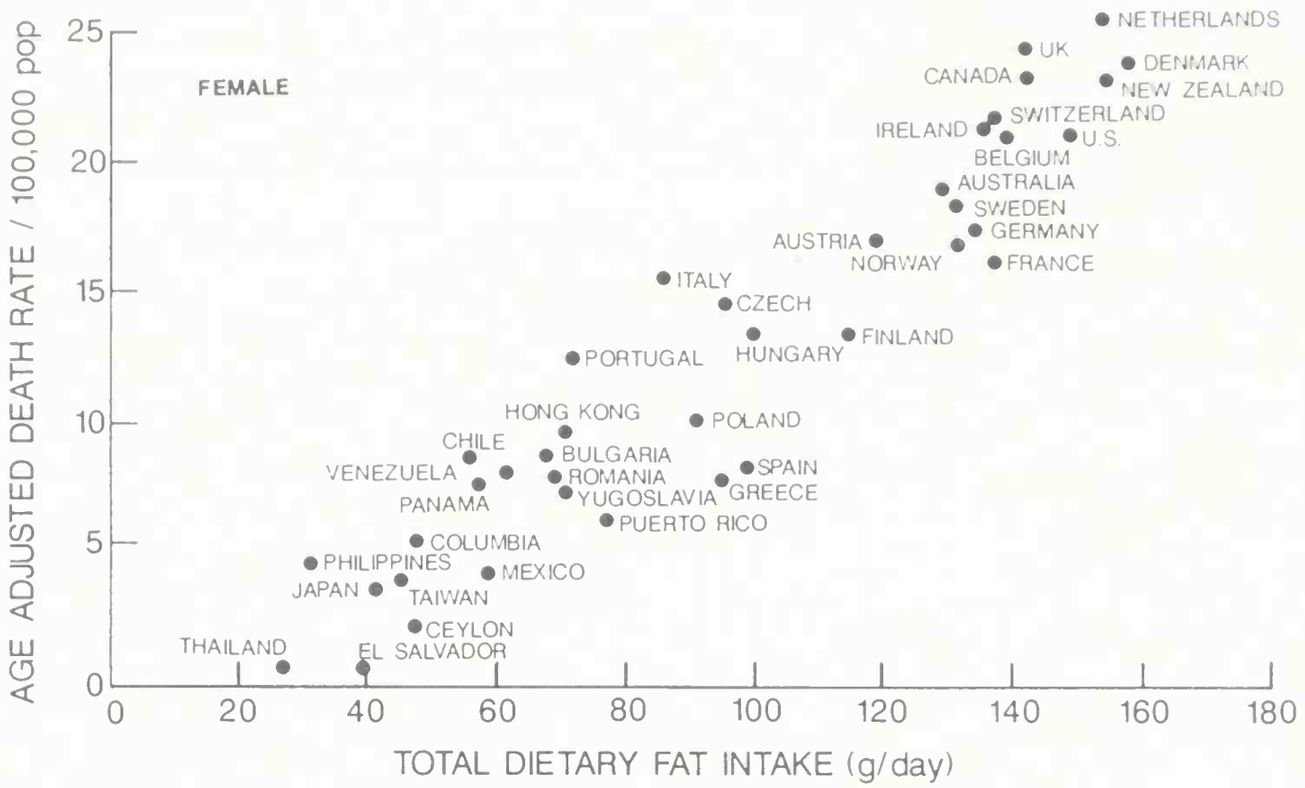


Figure 14–1. Relation of national per capita fat intake with risk of breast cancer mortality. (From Carroll, 1975; reproduced with permission.)

diet be reduced from an average of 40 percent to an average of 30 percent of calories, largely based on an anticipated reduction in breast cancer rates (Committee on Diet, Nutrition and Cancer, 1982). This has subsequently become the focus of a major health promotion campaign by the National Cancer Institute (National Cancer Institute, 1984). In this instance, epidemiologic findings were translated into public policy with remarkable speed. For purposes of illustration, the epidemiologic data addressing this relationship are examined in some detail.

ECOLOGIC STUDIES

The extremely strong international correlation between national per capita total fat "disappearance" and breast cancer rates has been noted; this correlation is primarily due to animal fat ($r = 0.83$), as the correlation for vegetable fat is considerably lower ($r = 0.18$) (Hems, 1978). The limitations of such data have been discussed earlier. In this instance, the association is potentially confounded by lean body mass, obesity, sedentary life-style, and by any of the many correlates of economic development, as the high incidence countries tend to be Westernized populations, whereas the low incidence countries tend to be nonindustrialized. Indeed, the correlation between gross national product and breast cancer mortality rate is 0.72 (Armstrong and Doll, 1975). Although Japan may appear to be an exception, because it is an industrialized country with low rates of disease, until recently its population largely consisted of peasant farmers; even in 1950 more than half of its population was rural (Nukada, 1975). Prentice and colleagues (1988) have examined the relationship between fat disappearance and breast cancer rates for 21 countries and were not able to explain the positive correlation on the basis of standard risk factors.

An examination of the scale of fat consumption in the data of Carroll reveals some of the problems of data quality using this approach (Fig. 14–1). The fat consumption estimate for the United States is 150 g/day; if this represents 40 percent of calories, then total energy intake would be about 3400 kcal per individual, including men, women, and children. A more reasonable estimate of true intake would be 2600 kcal for men and 1800 kcal for women (Beaton et al., 1979). To the extent that wastage of fat is greater in the wealthy, industrialized countries, these international correlations may, in part, represent a relationship between food wastage and breast cancer risk.

Other geographic correlations between fat intake and breast cancer risk are less striking. A positive correlation between per capita consumption of dairy fat and breast cancer rates has been noted for geographic areas within England, but fat intake from other sources was inversely related to breast cancer rates (Stocks, 1970). Within the United States, regional consumption of milk, an important fat source, is positively associated with rates of breast cancer, although consumption of eggs, a major cholesterol determinant, is inversely related to breast cancer rates (Gaskill et al., 1979).

MIGRANT STUDIES

Migrant studies have demonstrated that the large differences in breast cancer rates among countries are not attributable to genetic factors. Buell (1973) observed that the offspring of immigrants from Japan to the United States, but not the immigrants themselves, have breast cancer rates that are similar to those of the general American population. Polish women, however, who migrate to the United Kingdom or the United States (Staszewski and Haenszel, 1965; Adelstein et al., 1979), and Italian women who migrate to Australia (McMichael and Giles, 1988), themselves attain rates of breast cancer that are similar to the higher rates among women born in these countries, suggesting that the delayed effect among Japanese–Americans may be due to a slower acculturation process. This distinction is potentially important for dietary studies, because an exposure that acts only in childhood but is manifested decades later will be difficult to investigate. Although the potential period of susceptibility to dietary fat remains uncertain, large bowel cancer rates increase more rapidly than breast cancer rates within the same Japanese migrant populations (Haenszel and Kurihara, 1968), suggesting that relationships with adult diet are more likely to be seen for colon cancer.

SPECIAL POPULATIONS

The rates of breast cancer among special populations who have been consuming unique diets for long periods are of interest as an influence of diet should not be missed due to error in the measurement of individual diets or a limited follow-up period in a cohort study. It was originally reported that, compared with the general U.S. population, breast cancer mortality rates were lower among Sev-

enth Day Adventists, who consume relatively small amounts of meat and other animal products (Phillips, 1975). It was later appreciated, however, that these differences in mortality rates could largely be attributed to confounding by socioeconomic status. Although breast cancer incidence is positively associated with economic status in many countries, breast cancer mortality rates in this instance were inversely related to socioeconomic status presumably due to earlier diagnosis and more effective treatment. When compared with a U.S. white population of similar socioeconomic level, breast cancer mortality rates among a large group of Seventh Day Adventists were only slightly and nonsignificantly lower than expected (SMR = 0.85, Phillips et al., 1980). In striking contrast, the colon cancer mortality rate among the Seventh Day Adventists was only about half that of the comparison population. Although total fat intake of Seventh Day Adventists (36% of energy) is only slightly lower than the general U.S. population, vegetable fat is largely substituted for animal fat (Mills et al., 1988); thus these data do not support the hypotheses that animal fat or meat intake are specifically related to breast cancer.

Kinlen (1982) compared rates of breast cancer among orders of nuns who were either vegetarians or ate only small amounts of meat with rates among single British women (thus presumably controlling for parity); no significant differences were found. Because these women had typically entered their convent before age 20, a contrast in diet had existed for many decades. Although the conclusions of this study were limited by the modest number of cases of breast cancer among the nuns (62) and the limited quantitative data on their diets, this study illustrates the potential value of special exposure groups to evaluate dietary hypotheses. We cannot be certain that an unknown risk factor among the vegetarian nuns compensated for a protective effect of their diet, but this would require a complex hypothesis. If the rates of breast cancer had been different, the interpretation would be more difficult because many factors other than fat intake, both dietary and nondietary, differed between the groups of women.

The very strong correlation between dietary fat intake and breast cancer rates among five ethnic groups living in Hawaii (Kolonel et al., 1981; Fig. 14–2) has been interpreted by some as evidence of an etiologic relationship. An examination of the scale of this figure, however, indicates that a 34 percent increase in fat is associated with a 200 percent increase in breast cancer incidence. Although it is possible that dietary fat intake explains some of the differences in rates, the implausibly strong relationship, which is much stronger than suggested by the international correlations, increases the likelihood that the association observed in this study is confounded by some other factors.

SECULAR TRENDS

Major changes in breast cancer rates within one country provide further strong evidence that nongenetic factors have an important influence on the occurrence of this disease. For example, dramatic changes in breast cancer incidence have occurred in Iceland during this century (Bjarnason et al., 1974). This increase

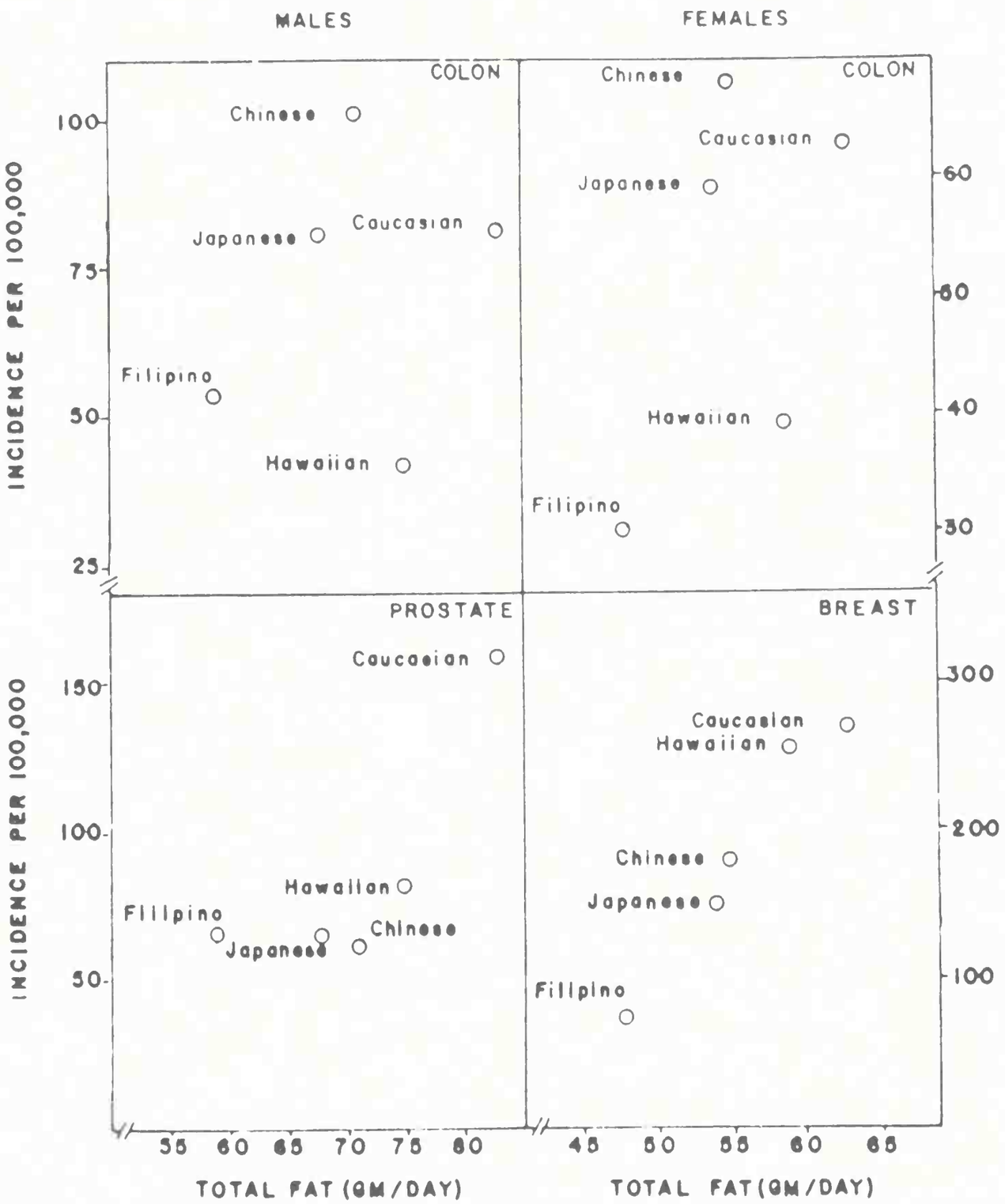


Figure 14–2. Dietary fat intake of five ethnic groups in Hawaii in relation to age-adjusted breast cancer incidence among women 45 years of age or older. (From Kolonel et al., 1981; reproduced with permission.)

was primarily in women 45 years of age and older (Fig. 14–3), which provides evidence to support the hypothesis of de Waard and co-workers (1964) that environmental factors differentially affect premenopausal and postmenopausal disease. Because the diet of the Icelandic population changed substantially over that period of time, becoming high in fat composition like other Western countries, these data are consistent with the hypothesis that fat intake causes breast cancer, but do not exclude other possible explanations, including an increase in total energy availability in relation to requirements.

In Japan, fat intake has increased dramatically since World War II; between 1955 and 1973 per capita fat intake increased more than 150 percent, whereas

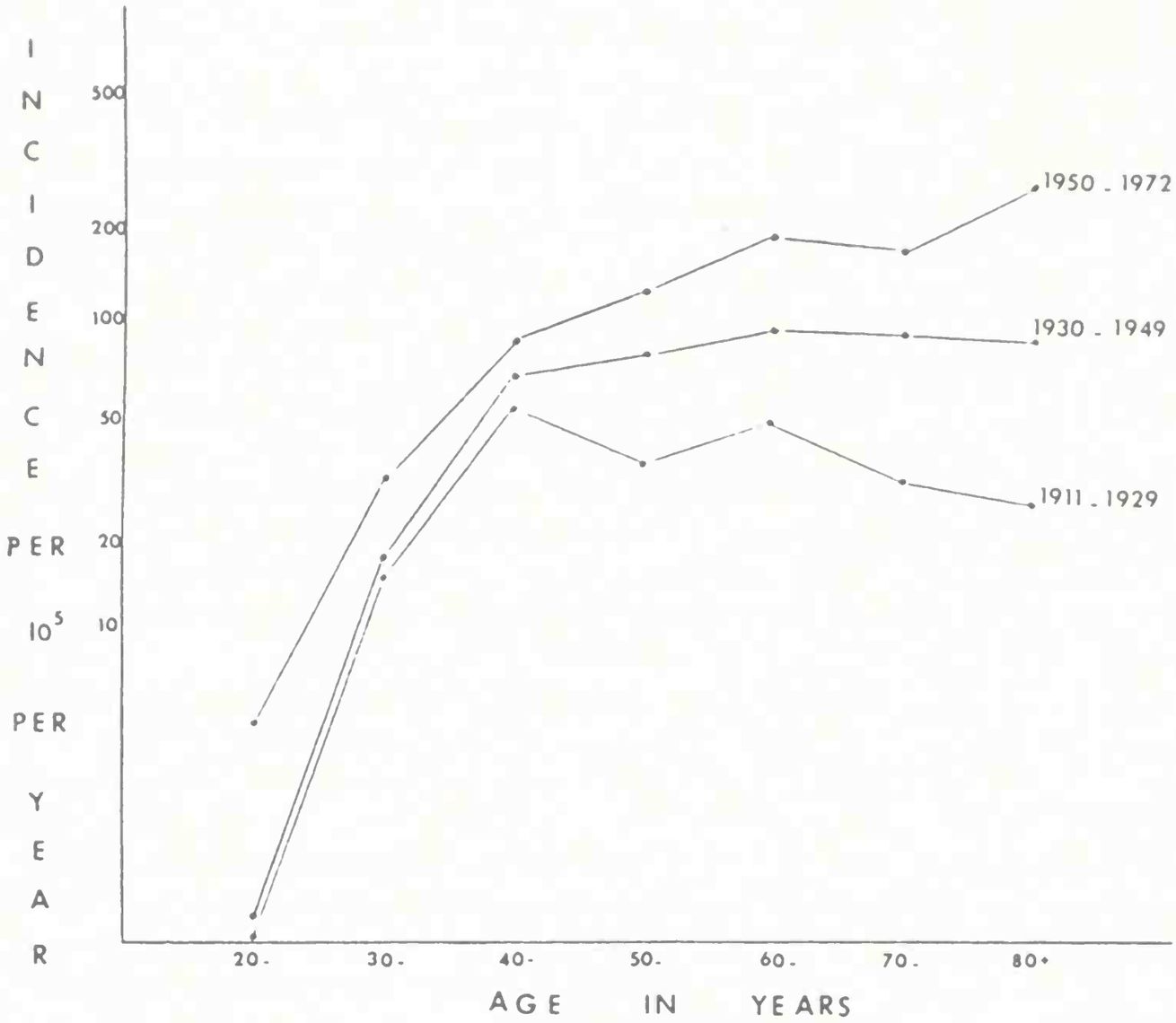


Figure 14–3. Age-specific incidence of breast cancer in Iceland for the three time periods 1911–1929, 1930–1949, and 1950–1972. (From Bjarnason et al., 1974; reproduced with permission.)

total energy intake remained unchanged (Fig. 14–4). During this period, age-adjusted breast cancer mortality rates increased much less markedly (Fig. 14–5); essentially no increase occurred among women 60 years of age and older (Hirayama, 1978; Fig. 14–6). Incidence rates of breast cancer also changed only slightly between 1963 and 1977 (Hanai and Fujimoto, 1982).

Possible explanations of the small changes in breast cancer rates in Japan, despite radical increases in fat intake, are that the diets of women changed less than men, or that diets of older women changed less than younger women. The striking increase in colon cancer mortality among Japanese women, which is also hypothesized to be related to diet and affects predominantly older individuals, however, argues against these explanations. As shown in Figure 14–5, age-adjusted colon cancer mortality through 1984 has increased almost in proportion to the increase in dietary fat between 1955 and 1973, and is now nearly equal to breast cancer mortality. None of these secular trends have been adjusted for changes in reproductive risk factors; it is possible that much of the small increase in breast cancer is due to later and fewer births. Reproductive patterns have changed dramatically in Japan; for example, between 1950 and 1975 the percentage of women having a first birth before age 19 decreased from 13.3 per-

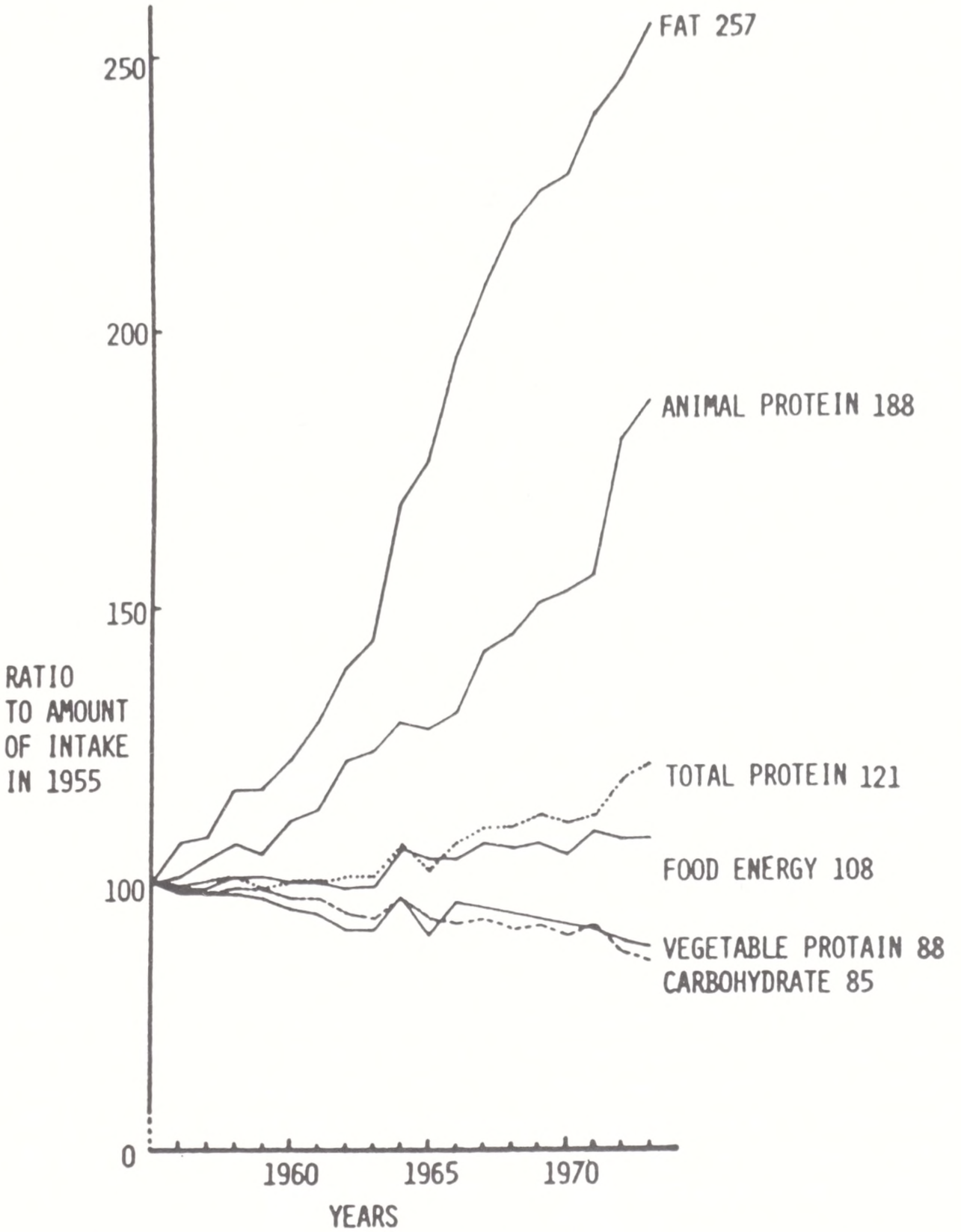


Figure 14–4. Change in amount of intake of selected nutrients in Japan from 1955 to 1973. (From Hirayama, 1978; reproduced with permission.)

cent to 4.1 percent (Health and Welfare Statistics Association, 1987). Future changes in breast cancer incidence are difficult to predict; if diet acts primary during childhood, the full effect of the dramatic nutritional changes in Japan may not be seen for several more decades.

Enig and colleagues (1978) have related consumption of different types of fats to the apparent increase in incidence of breast and other cancers in the United States. They have reported that the strongest association is that with the consumption of *trans*-fatty acids, which are fatty acids created in processes that

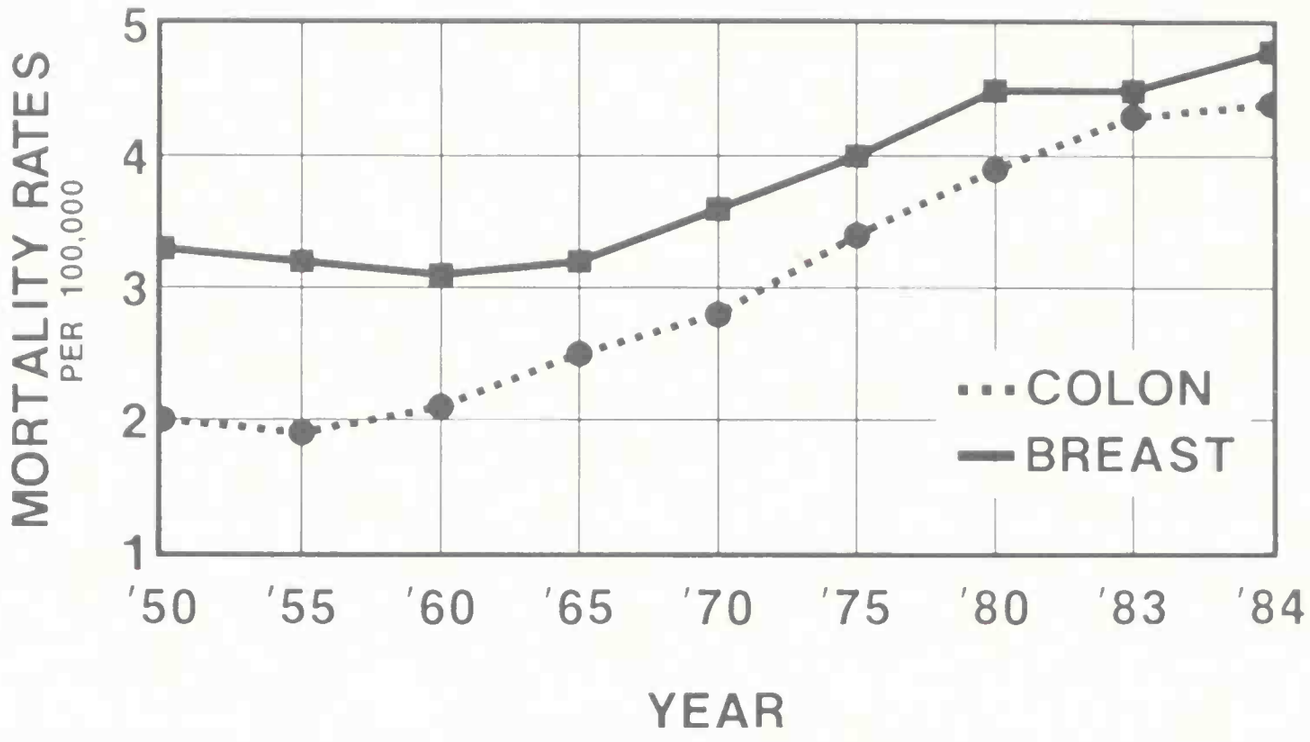


Figure 14–5. Age-adjusted breast and colon cancer mortality among Japanese women, 1950 to 1984. (Based on data from Health and Welfare Statistics Association, 1986.)

convert liquid vegetable oils to margarine and solid vegetable shortening. It is not clear, however, whether the U.S. incidence of breast cancer has actually been increasing beyond what would be expected on the basis of a trend toward later and fewer pregnancies.

Famines or other sudden changes in national diets due to war and social upheavals may potentially be useful to examine the latent period between change in diet and change in breast cancer rates. Ingram (1981) used this approach to examine the breast cancer mortality in England and Wales in relation to the marked changes in diet that occurred during World War II. He found positive correlations between breast cancer mortality and intake of meat, fat, and sugar, with maximal associations for a lag interval of 12 years. The marked differences in breast cancer mortality observed by Ingram, however, appear to be an artifact due to a change in procedures for coding deaths (Key et al., 1987). Thus it appears that the changes in diet during and after World War II in England had no important effect on breast cancer mortality.

CASE-CONTROL STUDIES

The relationship between fat intake and risk of breast cancer has been examined in relatively few case-control studies. In what has been described (Greenwald et al., 1987) as the most thorough case-control study, even 9 years after its publication, Miller and colleagues (1978) compared the diets of 400 Canadian women with breast cancer with those of 400 neighborhood controls. (The abstract of this paper was quoted earlier in this chapter.) Because of the exceptional quality of the data and the prominence this study has received as support for the dietary fat and breast cancer hypothesis, the findings are examined in detail.

In the study conducted by Miller and colleagues, three methods were used to measure dietary intake. A 24-hour recall was administered to familiarize subjects with the process of reporting food intake; a 4-day dietary record was completed to serve as a standard for validating their diet history questionnaire; and the dietary history questionnaire was completed by an interview that referred to a 2-month period 6 months earlier, with the intent of avoiding any influence of the cancer diagnosis or treatment on the report of diet. In their companion paper (Morgan et al., 1978), the authors state that they "believe that both the 24-hour recall and the 4-day diary are inherently less reliable for individual estimates of usual intake, and that when such estimates are required, the diet history should be used." They further appropriately note that in a case-control study there is reason to be concerned that the diagnosis or treatment of cancer could alter diet or its recall; this is additional reason to believe that data based on a 24-hour recall or a diet record collected after the diagnosis of cancer will not be relevant to disease etiology.

Although Miller and colleagues presented their data in several ways, for pur-

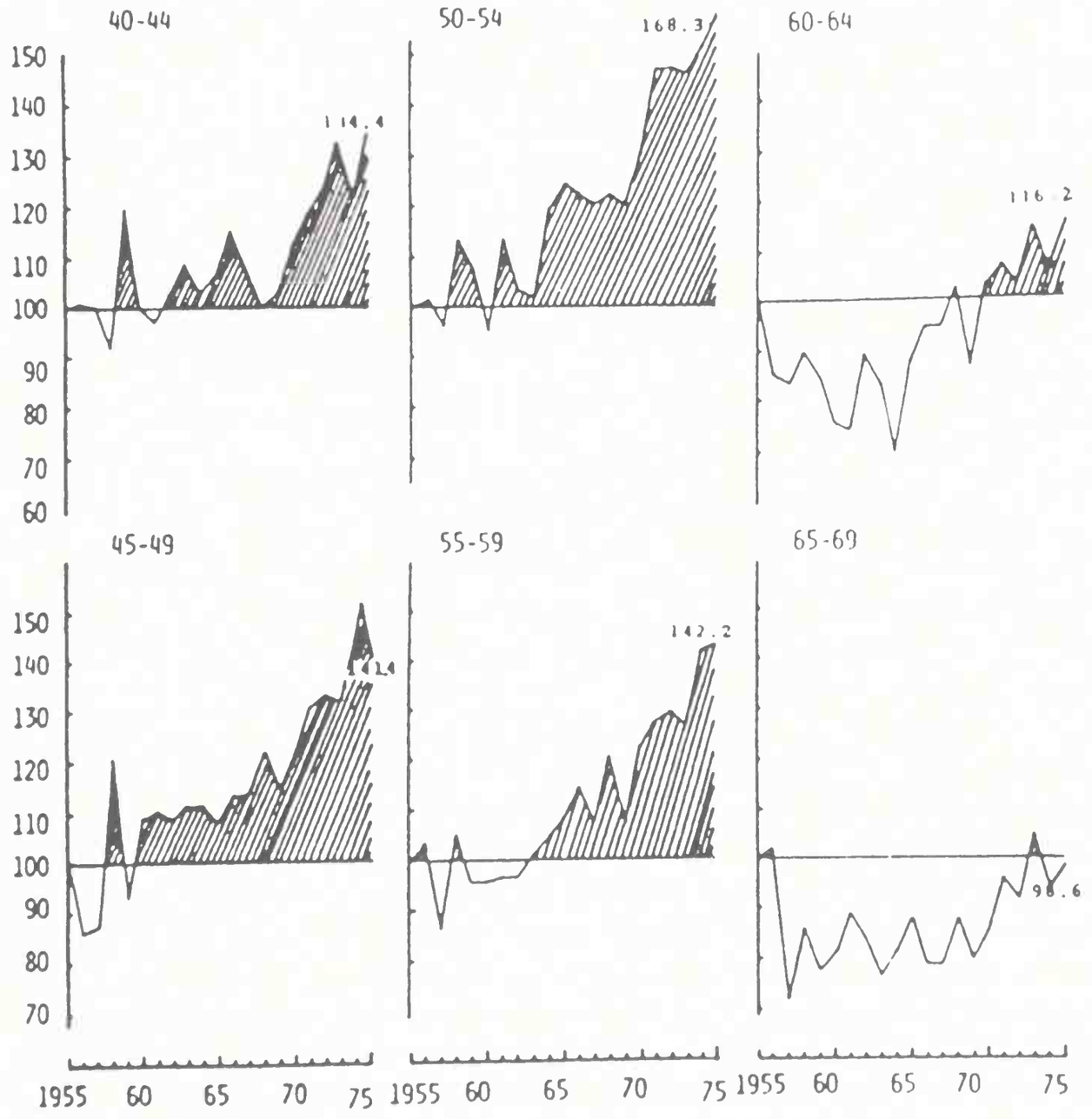


Figure 14–6. Age-specific death rates from breast cancer among Japanese women. (From Hirayama, 1978; reproduced with permission.)

Table 14–1. Mean nutrient intake based on 24-hour recall, 4-day diet record, and diet history questionnaire for breast cancer cases and controls, premenopausal and postmenopausal women combined

	Dietary method		
Nutrients and series	24-hour recall	4-day record	History
Total calories			
Cases	1697	1805	2280
Controls	1587	1785	2230
Difference[a]	109	20	50[d]
(n)[b]	(396)	(309)	(388)
	(p = 0.01)[c]		
Total fat (g)			
Cases	73.2	81.9	99.1
Controls	69.4	80.0	96.6
Difference	3.7	1.9	2.5
(n)	(386)	(309)	(395)
Saturated fat (g)			
Cases	28.0	32.0	38.3
Controls	26.2	30.3	37.1
Difference	1.8	1.7	1.2
(n)	(383)	(306)	(397)

[a]Mean difference for matched pairs.

[b]Number of pairs.

[c]One-sided p values, other comparisons not statistically significant.

[d]Difference erroneously given as 150 in original table.

Data from Miller et al., 1978.

poses of illustration the means for cases and controls are shown here. In epidemiologic studies we are ultimately concerned with the change in rates of disease as a function of change in exposure, in this case dietary fats. For this reason, epidemiologists typically divide continuous exposures into categories, use one of these categories as a reference, and compare rates or relative risks of disease for the other categories to those for the reference. Analyzed in this way, it frequently becomes apparent that substantial and important effects of the exposure are present even when the differences in mean values for cases and controls seem trivial (see Chapter 3). Although a thorough analysis will include a display of relative risks according to exposure level, it is useful to recognize that these data are derived from the continuous distributions for cases and controls and that comparisons of mean values for the two groups, given that they are approximately normal and have similar variances, can provide a concise and informative summary of the data as well as the most powerful statistical test of a difference between groups. Moreover, the potential for any systematic error or bias between cases and controls to distort the association can often be best appreciated by comparing means as the difference between their values is frequently within the range of highly plausible degrees of bias.

The primary findings from the study of Miller and colleagues are displayed in Table 14–1; although the authors provided information for other dietary lipid fractions, the data for calories, total fat, and saturated fat form the basis of their conclusions. It is apparent that the only statistically significant results are for

data based on the 24-hour recall, which the authors appropriately believe is likely to be both unreliable (subject to major random error) as well as biased in the context of this case-control study. For their preferred method of assessment, the dietary history questionnaire, the differences between cases and controls are small and not statistically significant, even by their one-sided test. Moreover, the magnitude of the difference in total or saturated fat intake between cases and controls is almost exactly proportional to that of total caloric intake. From these data it can be calculated that the cases reported 39.1 percent of their calories from fat and the controls 39.0 percent of calories from fat. For saturated fat, the values are 15.1 percent for cases and 15.0 percent for controls. (As discussed in Chapter 11, this is not the optimal method to adjust for total caloric intake in an epidemiologic context. However, for variables that are highly correlated with total energy intake the result will usually be very close to that obtained with regression analysis (Shekelle and Nichamen, 1987; Willett and Stampfer, 1987d. The optimal analysis would require the original raw data.) Thus, this study actually provides no support for the hypothesis that the fat composition of the diet affects risk of breast cancer; furthermore, it illustrates the potential for bias when collecting information about diet after rather than before the diagnosis of disease.

Although the overall data of the study by Miller and co-workers do not support the hypothesis that dietary fat increases the risk of breast cancer, it is possible that a positive association exists within a subgroup that is obscured by only examining the total group. Many believe that diet would have the strongest relationships with breast cancer among postmenopausal rather than premenopausal women because the international differences and secular changes within some countries are substantially greater for older rather than younger women. Miller and colleagues, therefore, examined women separately according to menopausal status (Table 14–2). Although none of the differences are statistically significant,

Table 14–2. Mean nutrient intake based on the diet history questionnaire for breast cancer cases and controls according to menopausal status

Nutrient and series	Premenopausal women	Postmenopausal women	Women aged 70 or more years
Total calories			
Cases	2373	2170	1614
Controls	2339	2115	1996
Difference	34	55	−383
(n)	(85)	(210)	(13)
Total fat			
Cases	109.4	91.8	58.9
Controls	104.2	90.7	71.7
Difference	5.2	1.1	−12.8
(n)	(88)	(213)	(13)
Saturated fat			
Cases	42.0	34.3	21.9
Controls	40.6	34.5	25.2
Difference	1.4	0.8	−3.4
(n)	(88)	(214)	(13)

Data from Miller et al., 1978.

the small overall differences in absolute total and saturated fat intake are almost entirely attributable to premenopausal women, in contrast with expectation. For postmenopausal women over the age of 70, the cases actually reported substantially less fat intake than controls (see Table 4 of the original article). For fat as a percentage of calories, the direction of the relationship among postmenopausal women was slightly opposite to that hypothesized; for total fat, cases reported 38.1 percent and controls 38.6 percent, whereas for saturated fat cases reported 14.6 percent and controls 14.7 percent. The authors of this report also provided relative risks for different levels of fat intake. Although some of the relative risks were greater than 1.0 (1.6 for total fat and 1.4 for saturated fat, comparing highest and lowest thirds), none approached statistical significance and no suggestions of any dose-response relationships were observed. Moreover, none was adjusted for total caloric intake, which would have reduced the observed associations.

As Miller and colleagues indicated, their dietary history questionnaire, like any other, was less than perfect; this would tend to underestimate or obscure true positive associations. The correlations between their questionnaire and a 4-day dietary record administered in close temporal proximity were 0.27 for calories, 0.31 for total fat, and 0.34 for saturated fat (Morgan et al., 1978). Although these correlations are rather low and could easily account for the failure to observe associations, they probably underestimate the validity of their questionnaire as a 4-day record itself does not represent a good estimate of usual intake.

In summary, this carefully conducted study does not support the hypothesis that the fat composition of the diet is associated with breast cancer incidence. It is reasonable to suggest that observed associations are likely to have been underestimated due to an imperfect measure of exposure. It seems inappropriate to conclude, however, that the use of an imperfect measure of diet means that a true association exists when no association, or one that is readily compatible with chance, is observed (see excerpt from abstract at the beginning of this chapter).

Howe (1985) has published an important methodologic paper on the use of dual response measures to improve the estimates of relative risks when both measures are imperfect. Unfortunately, the data used to illustrate his methods were the 24-hour recall and diet history information from the study by Miller and colleagues described earlier. The very strong imputed relative risks for saturated fat obtained by Howe (5.9 for high versus low intake) has frequently been cited as support for the fat and breast cancer hypothesis. It will be readily appreciated that these strong relative risks are largely due to the association based on the 24-hour recall, which the original authors appropriately recognized as artifactual. Sophisticated statistical manipulation cannot correct for primary data that are inherently biased.

The largest case-control study to date of dietary fat intake and breast cancer incidence was reported by Graham and colleagues (1982). Fat intake, estimated with a simple food-frequency questionnaire, among 2024 women with breast cancer was essentially identical to that reported by 1463 control women attending a hospital for a variety of benign conditions.

In a recent Australian study involving 451 case-control pairs, Rohan and

colleagues (1988) also found no evidence of a positive association between total fat intake and breast cancer risk [relative risk = 0.90 (95% confidence interval = 0.59 to 1.38) for the highest compared with the lowest quintile]. Similarly, Hirohata and colleagues (1987) found only weak and nonsignificant associations between intake of total and specific types of fat in a case-control study among 183 Japanese and 161 white case-control pairs living in Hawaii. In this study the quite limited list of foods was assessed using a 1-week recall; these did not allow a calculation of total energy intake.

Studies conducted among the general populations of these and other Western countries share the constraint that few women consume a diet with less than 30 percent of calories from fat. For this reason the Japanese case-control study of Hirohata and colleagues (1985) conducted among 212 women with breast cancer and an equal number of each of hospitalized and neighborhood controls is of special interest. In this study, the mean daily total fat intake reported by cases was 51 g, by hospital controls was 52 g, and by neighborhood controls was 52 g. A similar lack of any substantial difference was seen for both animal fat and vegetable fat. A similar absence of association between fat intake and breast cancer risk was observed in a small case-control study conducted in Greece (Katsouyanni et al., 1988).

Although the published case-control studies of fat intake and breast cancer provide little support for an etiologic relationship, the degree to which they are informatively null is limited by lack of information on the true between-person variation in fat intake in the study populations and, with the exception of the Miller study, the validity of the questionnaire employed. The diet record data from the validation study imbedded in the investigation of Miller and colleagues could be subjected to analysis of variance to determine the within-person and between-person variation and thus an improved estimate of questionnaire validity (see Chapter 12), but this apparently has not been done.

A number of other case-control studies of breast cancer have contained assorted questions about the use of specific high-fat foods, but were not extensive enough to provide an estimate of total fat intake. One of these studies (Lubin et al., 1981) has frequently been cited as supporting the fat and breast cancer hypothesis. Five hundred and seventy-seven women hospitalized with breast cancer were asked about their frequency of use of eight foods. Apparently as an after-thought, it was decided to obtain similar information from a series of control women. This required hiring a new interviewing staff, which then administered the questionnaire at home to the 72 percent of the general population sample willing to participate. The associations found for the use of beef, pork, and sweet desserts were striking (Table 14–3). If these relationships were truly this strong, one would expect Seventh Day Adventists or vegetarian nuns to have markedly lower rates of breast cancer and associations for these foods to be obviously apparent in the case-control studies of Graham and co-workers and Miller and co-workers and the prospective data described later of Phillips and Snowdon (1983) and Willett and colleagues (1987a). In retrospect, it seems unlikely that these associations can be correct and that they are much more likely to be artifacts of the noncomparable manner in which the data were collected. The findings of this study suggest that adherence to the basic epidemio-

Table 14–3. Age-adjusted relative risk of breast cancer for various food items categorized by tertiles

	Level[a]	Cases	Controls	RR	95% CI
Beef	6	87	127	1.53[b]	(1.1, 2.1)
	5	274	301	2.25	(1.8, 2.9)
	1-4	197	397	1.00	
Pork	4-6	320	398	2.16[b]	(1.6, 2.9)
	3	120	181	1.76	(1.3, 2.5)
	1-2	112	246	1.00	
Fowl	4-6	368	621	0.87	(0.6, 1.4)
	3	151	151	1.54	(0.9, 2.5)
	1-2	39	53	1.00	
Fish	4-6	288	438	1.02	(0.8, 1.3)
	3	141	185	1.26	(0.9, 1.7)
	1-2	129	201	1.00	
Eggs	6-5	160	254	0.84	(0.6, 1.2)
	4	293	449	0.88	(0.6, 1.2)
	1-3	105	121	1.00	
Cheese	6	199	310	1.11	(0.9, 1.4)
	5	126	159	1.37	(1.0, 1.9)
	1-4	232	354	1.00	
Creams (full, sour, ice, whipped)	5-6	79	120	0.92	(0.7, 1.2)
	4	184	307	0.90	(0.7, 1.2)
	1-3	290	301	1.00	
Sweet desserts	5-6	183	224	1.45[c]	(1.1, 1.9)
	4	189	286	1.26	(1.0, 1.6)
	1-3	176	316	1.00	

[a]Food frequency levels are defined as: 6, daily; 5, 4-6 days/week; 4, 1-3 days week; 3, >1 day/month and <1 day/week; 2, ≤1 day/month; 1, never.

[b]Test for linear trend, $p < 0.001$.

[c]Test for linear trend, $p = 0.01$.

From Lubin et al., 1981.

logic principles of comparability in data collection procedures for cases and controls is not merely a hypothetical issue.

In other case-control studies that included a limited list of foods, Phillips (1975) found a significant association between fried potatoes and the risk of breast cancer in a Seventh Day Adventist population, but apparently found no association with meat. In an Italian case-control study, Talamini and co-workers (1984) reported a positive association with intake of milk and dairy products, but not meat. In a French case-control study, Le and co-workers (1986) found positive associations with the use of cheese and full cream milk, but not with the use of butter or yogurt. In a large case-control study conducted in western Canada, Hislop and co-workers (1986) found positive associations with the use of gravy, beef, and pork among premenopausal women; among postmenopausal women, the risk of breast cancer was associated only with consumption of pork. In addition to the limitations of the case-control studies that had a reasonably comprehensive assessment of fat intake, these studies do not allow any examination of whether the sporadically observed associations are confounded by total energy intake.

In two other case-control studies, a reasonably comprehensive dietary ques-

tionnaire was employed but the data were not analyzed by computing fat intake. Katsouyanni and colleagues (1986), in a Greek study consisting of 120 cases and 120 hospitalized controls, combined individual foods into food groups, which provides a practically informative result. For both food groups that contribute substantially to fat intake, "meat, fish, eggs" and "dairy products," nonsignificant inverse relationships were seen with risk of breast cancer. These findings were in striking contrast to the strong positive association between the meat food group and risk of colon cancer found in another case-control study conducted by the same group using the same questionnaire and overall design (Manousos et al., 1983). In an Israeli study that used a 250-item dietary questionnaire, Lubin and co-workers (1986) compared the diets of 818 breast cancer patients with those of both a surgical and neighborhood control series. The analysis was conducted by adding the frequencies of all foods containing more than 20 percent fat (whether this is on the basis of weight or energy is not clear) and divided the sum of these frequencies into four categories. Positive trends with increasing intake of fat foods were seen using both control groups, but only for surgical controls 50 years of age or older was the trend statistically significant. Unfortunately, this method of analysis does not make maximal use of the data as the frequencies of use are simply added without giving any weight to the amount of fat in the various foods that contain more than 20 percent fat. Thus, the consumption of olives would count the same as the consumption of beef steak.

PROSPECTIVE STUDIES

The potential for biased associations due to selective participation or differential recall of past diet inherent in case-control studies is eliminated in prospective studies of the fat and breast cancer relationship. Thus far relatively few prospective studies have been published, which is largely a function of the cost and time involved in conducting such studies. In a prospective study from Japan based on a very limited number of dietary questions, Hirayama (1978) reported a higher incidence of breast cancer among women consuming meat daily; however, the total number of cases among women who ate meat daily was only 14, which precluded any detailed analysis.

Phillips and Snowdon (1983) examined the relationship between meat consumption and mortality due to breast cancer during a 21-year follow-up period of California Seventh Day Adventists. During this period 186 women died of this disease; the breast cancer mortality rates (per 100,000 person-years) were 47.8 for no use of meat, 58.3 for meat use one to three times per week, and 56.9 for use of meat four or more times per week (p for trend = 0.28). These data are of particular value because of the large portion of the population that consumed no meat at all. This Seventh Day Adventist population also provided a rare opportunity to evaluate the effects on breast cancer risk of dietary changes at various ages because the age at adopting a vegetarian life-style can usually be determined. Mills and colleagues (1988) found that, in addition to the lack of association with meat intake, cheese, milk, and eggs were not related to risk of

death due to breast cancer. Moreover, among women who did not eat meat, those who adopted a vegetarian life-style earlier in life tended to have a higher, rather than lower, risk of breast cancer.

The largest prospective study reported to date to include a calculation of total fat intake was based on a dietary questionnaire (see Appendix of Chapter 5) completed by 89,538 registered nurses aged 34 to 59 years in 1980 (Willett et al., 1987a). During 4 years of follow-up, 601 cases of breast cancer were diagnosed among the participants. After adjustment for known determinants of breast cancer, the relative risk of breast cancer among women in the highest quintile of calorie-adjusted total fat intake, as compared with women in the lowest quintile, was 0.82 (95% confidence interval, 0.64 to 1.05), and for saturated fat intake the corresponding relative risk was 0.84 (0.66 to 1.08, Table 14–4). Similar nonsignificant inverse trends were seen for calorie-adjusted linoleic acid and cholesterol and for the same dietary lipids not adjusted for caloric intake. Furthermore, no evidence of a positive association with meat intake was observed (Table 14–5).

This study was unique because it included a validation study that provided a measurement of the distribution of fat intake in the study population indepen-

Table 14–4. Age-adjusted relative risk (RR) of breast cancer according to quintile of calorie-adjusted intake of total and saturated fat, linoleic acid, and cholesterol

	Quintile for intake					χ_1 trend
Measurement	(Low) 1	2	3	4	(High) 5	(p value)
Total fat						
No. of cases	145	112	122	110	112	
No. of women	17,841	17,909	17,924	17,929	17,935	−1.57
Multivariate RR[a]	1.0	0.80	0.88	0.80	0.82	(0.11)
(95% confidence limits)	—	(0.62,1.02)	(0.69–1.12)	(0.63,1.03)	(0.64,1.05)	
Saturated fat						
No. of cases	146	112	126	105	112	
No. of women	17,848	17,910	17,938	17,915	17,927	−1.86
Multivariate RR[a]	1.0	0.80	0.91	0.77	0.84	(0.06)
(95% confidence limits)	—	(0.63,1.03)	(0.72,1.16)	(0.60,1.00)	(0.66,1.08)	
Linoleic acid						
No. of cases	151	119	103	115	113	
No. of women	17,848	17,875	17,909	17,961	17,945	−1.42
Multivariate RR[a]	1.0	0.84	0.75	0.86	0.88	(0.16)
(95% confidence limits)	—	(0.65,1.07)	(0.58,0.97)	(0.67,1.10)	(0.69,1.12)	
Cholesterol						
No. of cases	118	129	119	129	106	
No. of women	17,916	17,935	17,878	17,920	17,889	−0.76
Multivariate RR[a]	1.0	1.06	1.02	1.07	0.91	(0.43)
(95% confidence limits)	—	(0.82,1.38)	(0.79,1.32)	(0.83,1.38)	(0.70,1.18)	

[a]The model includes indicator variables for quintiles 2 to 5 of fat intake, age (five categories), a maternal history of breast cancer, a sister with a history of breast cancer, nulliparity, age at first birth <23 years, current smoking, highest quintile for relative weight, history of benign breast disease, postmenopausal status, and alcohol consumption (three categories).

Table 14–5. Relation of meat-eating with incidence of breast cancer among 89,538 women during 4 years of follow-up[a]

	Frequency of eating				
	<1/wk	1/wk	2–4/wk	5–6/wk	Daily or more
Beef, pork, or lamb as a main dish					
Cases	67	186	258	45	33
Total women	11,416	27,301	36,453	8,820	5,305
Age-adjusted RR	1.00	1.21	1.26	0.90	1.05
	χ, trend = −0.24				
Beef, pork, or lamb as a sandwich or mixed dish					
Cases	52	141	227	134	31
Total women	8,885	22,903	30,234	22,609	4,920
Age-adjusted RR	1.00	0.98	1.25	0.98	1.00
	χ, trend = 0.27				

[a]Numbers add to slightly less than 89,538 due to missing data.
RR = relative risk.
Data based on study of Willett et al., 1987.

dent of the study questionnaire. Based on 28 days of diet records completed by 173 participants, the mean values for lowest and highest quintiles of absolute fat intake in this population were 47 and 98 g/day. Expressed as percentage of total caloric intake, the means for extreme quintiles were 32 and 44 percent. Ideally, it would be of considerable interest to examine the effect of fat intake below 30 percent of calories; however, these data suggest that this will be difficult in a contemporary general U.S. population as relatively few eat so little fat. Nevertheless, the degree of variation in fat intake with the study population is of interest as it corresponds closely to the current advice to decrease fat intake by one fourth from an average of 40 percent of calories to 30 percent of calories, and is, therefore, sufficient to evaluate the effect of these recommendations. For saturated fat intake, most highly suspected because of the international correlations, the variation in intake was greater; the means of lowest and highest quintiles for absolute intake were 16 and 35 g/day (a 117% increase from lowest to highest) and in relation to caloric intake were 11 and 17 percent (a 55% increase).

Even if adequate variation in dietary fat exists within the cohort, useful findings will be obtained only if the dietary questionnaire employed can discriminate among individuals. In this case, information on the performance of the questionnaire was provided by the validation study (Willett et al., 1985a and Chapter 6). Briefly, the correlation between the dietary questionnaire completed at the end of the year of diet record keeping and the average intake from the 28 days of diet records completed by each participant was 0.53 for calorie-adjusted total fat and 0.59 for calorie-adjusted saturated fat.

This degree of validity in measuring dietary fat is certainly far from perfect, but appears to be comparable to many measurements used in epidemiology, such as Quetelet's index (see Chapter 10). It is thus reasonable to consider how this degree of error could affect the findings. In general, error in the measurement of exposure that is random with respect to disease status has two impli-

cations. One is that the observed relative risk is closer to 1.0 than the true relative risk and the other is that the observed confidence intervals are narrower than the true confidence intervals (see Chapter 12). In this instance, where no significant association was observed, the focus of interest is in the upper bound of the confidence interval. In other words, what is the upper limit of the plausible relative risks that are compatible with the observed data from the study? It is also of interest that the observed relative risks in this instance were less than 1.0, in the opposite direction of the hypothesis. As pointed out by Potter (1987), a correction for measurement error will, therefore, move the relative risk further from unity; for total fat intake he estimated the corrected relative risk would be approximately 0.6 to 0.7. Statistical methods to correct the observed confidence intervals for measurement error (which includes uncertainty due to the measurment error itself as well as uncertainty in the estimation of the measurement error) are not yet published at this time. One method (Rosner et al., 1989), however, provides this information based on correction of logistic regression coefficients and their standard errors. Using this method and the data from the validation study, the observed relative risk (adjusted for age, alcohol intake, and calories) for the highest versus the lowest quintiles of saturated fat intake was 0.85 (95% confidence intervals, 0.67 to 1.07), which was corrected to 0.76 (95% confidence intervals, 0.50 to 1.13). Note that the point estimate moved away from one and the width of the confidence intervals increased. The upper bound was still only slightly above unity, indicating that, after accounting for error in the measure of fat intake, the data are compatible with only a very weak positive association. Even if the relative risk was centered on the observed relative risk, it is apparent that only weak positive associations would lie within the expanded confidence interval. Short of finding a significant inverse association, it is difficult for any study to exclude the possibility of a very small positive association. It is apparent, however, that the failure of this study to find a substantial positive association between fat intake and breast cancer incidence cannot simply be explained by imperfect measurement of exposure.

It is quite possible, even likely, that the latent period between exposure and disease was not represented in this study as the maximum follow-up time was 4 years, and that fat intake, therefore, might still influence breast cancer risk. This possibility cannot be eliminated as the latency period for breast cancer is unknown. In laboratory animals, however, dietary fat acts as a promoter, having an effect during the later stages of carcinogenesis (Hopkins et al., 1979). Moreover, even though the follow-up period was limited, dietary intake tends to be correlated over time (Rohan and Potter, 1984), so that the baseline assessment also reflected previous intake over a vaguely defined period of years. This study, however, does not address the possible influence of fat intake much earlier in life, such as during adolescence. Indeed, it is unclear whether the influence of diet composition in youth on risk of breast cancer can be effectively studied in adults as the recall as an adult of food intake during that period has not yet been demonstrated.

It should be noted that the age distribution of the Nurses' Health Study cohort was truncated so that the oldest participant was 59 years of age in 1980

when the dietary data were collected. Although there appeared to be no suggestion of any positive association between dietary lipid intake and risk of breast cancer in either premenopausal or postmenopausal women, an association among older postmenopausal women could not be excluded. There is no suggestion from any of the case-control studies, however, that a stronger positive association with dietary fat exists among older postmenopausal women; in the study of Miller and colleagues (1978), the direction of the relationship was actually inverse among this group.

The Nurses' Health Study illustrates another advantage of a prospective study as it is possible, with additional follow-up, to examine the relationship of the same dietary exposure data with risk of breast cancer at different latent periods. Thus, it will be of considerable interest to know the nature of this association after 10 or 15 years. In addition, repeated dietary questionnaires are being completed by participants at intervals of 2 or 4 years so that the effects of recent versus more remote diet can be distinguished with large numbers of endpoints. This information can also be used to examine the relationship between change in diet and change in risk of breast cancer, which has obvious public health implications.

The relationship of fat intake with incidence of breast cancer was also examined prospectively among 5485 women who completed a 24-hour recall as part of the First National Health and Nutrition Examination (Jones et al., 1987). During an average follow-up time of 10 years, 99 cases of breast cancer were diagnosed. Statistically significant inverse associations were seen for both total and saturated fat intake (for total fat, relative risk = 0.34, 95% confidence limits = 0.16 to 0.73 for the highest compared with the lowest category). After adjustment for total energy intake, this inverse association was somewhat weaker (relative risk = 0.66, 95% confidence interval = 0.33 to 1.31). The strong inverse associations seen in this study are particularly surprising as the dietary data were derived from a single 24-hour recall; the substantial within-person variation associated with this method would have made any realistic association difficult to detect. As suggested by the wide confidence intervals and the divergence from other findings, the very strong apparent protective effect of fat in this study is likely to reflect a major element of random error.

In summary, the case-control and cohort studies that have been sufficiently comprehensive to allow a computation of total fat intake have been remarkably consistent in failing to provide evidence of a positive association between the dietary lipid composition of the diet and risk of breast cancer. Sporadic associations have been observed between meat and dairy products in some case-control studies with a limited dietary assessment. These associations, however, have not been consistently observed in these limited studies or the more comprehensive studies, and are inconsistent with the rates of breast cancer in Seventh Day Adventists (Phillips et al., 1980) and vegetarian nuns (Kinlen, 1982). Moreover, if true, these positive associations with meat intake should have been readily observable in the prospective studies of Phillips and Snowdon (1983) and Willett and colleagues (1987a). It is, thus, most likely that these sporadic findings represent the play of chance combined with a tendency for positive findings within a study to be emphasized.

Biologic Plausibility

Some support for the hypothsis that high levels of dietary fat increase the rate of breast cancer is derived from studies that have related dietary factors to estrogen fractions, which are, in turn, thought to be related to breast cancer risk (Willett and MacMahon, 1984). Among postmenopausal women, omnivores on a high fat diet had higher urinary excretion of estriol and total estrogens (Armstrong et al., 1981) and higher plasma levels of estrone and estradiol (Goldin et al., 1981) than vegetarian women. The relatively low plasma levels of estrogen among postmenopausal vegetarians in the last-cited study were shown to be at least partly the result of greatly enhanced fecal excretion of estrogens. Feeding a high-fat, Western diet to postmenopausal black South African women who typically consumed a low-fat vegetarian diet caused an apparent decrease in levels of luteinizing hormone, follicle-stimulating hormone, and prolactin (Hill et al., 1980). In this study only a small increase in estradiol level was observed, and other estrogen fractions were apparently not measured. Hagerty and colleagues (1987) recently conducted a cross-over trial of a high fat diet (46% of calories) versus a low fat diet (25% of calories) among six women. No effect was seen on plasma or urinary levels of estrone, estradiol, or plasma progesterone or prolactin.

Among premenopausal women, inconsistent associations of diet and sex hormones have been observed. Studies among vegetarian and nonvegetarian Seventh Day Adventist teenagers (Gray et al., 1982a) and among teenage girls in four countries with large differences in breast cancer rates (Gray et al., 1982b) found no meaningful associations between plasma or urinary estrogen levels and dietary factors. Although premenopausal American women who were omnivores were found to have higher levels of plasma estrone and estradiol than vegetarians (Goldin et al., 1981), an apparent decrease in estradiol level was caused by feeding a high-fat, Western diet to premenopausal South African women (Hill et al., 1980). A decrease in fat intake from 35 to 21 percent of calories among women with cystic breast disease was associated with a reduction in estrone, estradiol, and estriol (Rose et al., 1987a), as well as prolactin (Rose et al., 1987b); however, total caloric intake among participants was also reduced by an average of 23 percent. The interpretation of these inconsistent findings is compounded by the difficulty in establishing a clear relationship between estrogen levels and risk of breast cancer (Buring et al., 1987).

In addition to the possibility of acting through an effect on estrogen metabolism, several other mechanisms have been proposed by which dietary fat may increase the risk of breast cancer. Wynder and colleagues (1976) have suggested that fat may affect breast cancer risk by altering prolaction secretion. The existence of such a mechanism is supported by some animal studies (Chan and Cohen, 1974), although it has not been established that prolactin secretion is related to breast cancer in humans. Another potential mechanism relates to the intake of polyunsaturated fats, which are subject to in vivo peroxidation. Highly reactive radicals generated in this process may damage DNA and other macromolecules, ultimately leading to neoplasia (Ames, 1983). It is of note, however, that the international correlations that underlie the fat and breast cancer hypoth-

esis are based on an association with animal fats, not vegetable fats, which are the primary source of polyunsaturated fatty acids in most countries. For example, Japanese diets differ from those in the United States in their saturated fat content, but not in the amount of polyunsaturated fatty acids (Insull et al., 1968).

It has also been suggested that dietary fat may be linked to breast cancer through increased caloric intake and the development of obesity. Adipose tissue can convert androstenedione to estrone (Grodin et al., 1973) and thus makes an important contribution to circulating levels in postmenopausal women. This would not be an effect of dietary fat specifically, however, as one can develop excess adipose on a diet that is not high in fat composition. Moreover, a clear positive association does not exist between obesity and incidence of breast cancer (discussed later).

Construction of a plausible mechanism whereby fat may affect the risk of breast cancer is seriously hindered by our basic ignorance of the pathophysiology of this disease. We do not really understand the steps that lead to breast cancer and lack an established measurable intermediary factor analogous to serum cholesterol or glucose in the case of coronary heart disease. Without an established biochemical or molecular precursor lesion, the search for a mechanism may be premature until it is reasonably established that, indeed, high fat diets increase the risk of human breast cancer.

A RE-EXAMINATION OF THE ANIMAL DATA

Although the hypothesis that high fat diets cause breast cancer in humans has largely been based on animal studies, the interpretation of these laboratory findings is controversial. This issue cannot be reviewed in detail here, but is discussed at length by Birt (1986) and in the proceedings of a recent symposium (Pariza and Simopoulos, 1987). A central question is whether fat intake has an effect on mammary cancer other than its contribution to total energy intake; this issue is remarkably parallel to that facing epidemiologists discussed in Chapter 11.

There is little question that restriction of energy intake dramatically lowers the incidence of mammary tumors (Tannenbaum and Silverstone, 1953). Because fat is uniquely dense in its energy content, a low fat diet is confounded by a reduction in energy intake unless strict care is undertaken to insure that the available energy intake is held constant. This issue has been studied by Boissonneault and co-workers (1986), who found, like many others, that rats on a low-fat ad lib diet had lower tumor incidence than those on a high-fat ad lib diet (Table 14–6). When the high fat diet was restricted so that total energy intake was about 20 percent lower than the ad libitum intake, however, the tumor incidence was reduced by 90 percent. This issue is further complicated by the finding that the net energy available to living organisms from macronutrients is not strictly proportional to the classic values obtained by bomb calorimetery (4 kcal/g of carbohydrate or protein and 9 kcal g of fat). Because more energy is required for the absorption and metabolic processing of carbohydrate and protein, in at

Table 14–6. Mammary tumor incidence in rats fed a diet with different fat contents

	Dietary regimen		
	High-fat, ad lib	Low-fat, ad lib	High-fat, restricted
kcal consumption per day	41 kcal	42 kcal	34 kcal
Fat consumption per day	2.7 g	0.6 g	2.2 g
Body weight	217 g	190 g	182 g
Body composition			
% body fat	24%	16%	25%
% body protein	20%	23%	20%
Retained energy	752 kcal	532 kcal	634 kcal
Tumor incidence	73%	43%	7%

From Boissonneau, H. et al., 1986.

least some situations they may provide approximately 20 percent less available energy than has been used in typical calculations (Donato, 1987).

Albanes (1987) has performed a meta-analysis of diet and mammary cancer experiments in mice conducted over the last 50 years. An extremely strong overall positive association was seen for total energy intake; however, after adjustment for total energy intake, the fat composition was actually weakly inversely related to incidence of mammary tumors. In another review of animal experiments, Birt (1986) concluded that evidence did exist for an effect of dietary fat independent of total energy intake. There is clearly not a consensus at present that the total fat composition of the diet influences the risk of mammary tumors; thus the existence of such a relationship cannot be used as an argument that a similar relationship should exist in humans.

A fundamental question related to the laboratory findings is whether any particular rodent model has relevance to human breast cancer. Ironically, this is a difficult issue to prove or disprove without firm human data. In toxicity studies, the analogy with humans is more direct as it involves fewer assumptions to presume the different species will respond similarly; even in this situation many exceptions can be found. In most studies of diet and cancer in animal models, however, one is examining the effect of diet on a cancer that is caused by an inducing agent that may be irrelevant to humans.

Using a quite different approach, Sonnenschein and colleagues (1987) have recently conducted a case-control study of breast cancer in dogs by interviewing owners about the usual foods consumed by their animals. Compared to human populations, an extremely wide intake in the fat composition of the diet was observed among these animals that varied little from day to day. Compared with controls having other cancers, and also a series of cancer-free controls, no association was observed with the fat composition of the diet.

Apart from possible effects of total fat intake on the occurrence of mammary tumors in animals, it has been suggested that the fatty acid composition of the diet has an independent relationship with this malignancy. In particular, it has been suggested that polyunsaturated fat may be particularly deleterious (Carroll and Hopkins, 1979; Hopkins and Carroll, 1979). This appears to be inconsistent with human data based on international correlations, case-control studies

(Miller et al., 1978), and prospective analyses (Willett et al., 1987a). Ip (1987) has suggested that little relationship exists in rodents once the essential requirements for linoleic acid have been met. Karmali (1987) has reported that high intake of omega-3 fatty acids, primarily obtained from marine oils, actually inhibits the occurrence of mammary tumors.

ALTERNATIVE HYPOTHESES

The failure of most case-control and cohort studies to confirm the hypothesis that a diet high in total or saturated fat composition increases the incidence of human breast cancer leaves the large differences in breast cancer rates among countries unexplained. Many alternative hypotheses exist, including the differences in intake of selenium and other minerals (Schrauzer et al., 1977), marine oils (Karmali, 1987), alcohol (Willett et al., 1987b; Schatzkin et al., 1987; Longnecker et al., 1988), and specific vegetables (Knox, 1977; Kamiyama and Michioka, 1983). Although a combination of these factors may contribute to differences in rates between countries, one alternative explanation for the large differences is that the unequivocal protective effect of energy restriction found in animal studies also applies to humans. More specifically, this hypothesis suggests that energy intake sufficiently restricted in relation to requirements during development so as to reduce adult size will reduce the incidence of breast cancer in humans. This hypothesis has been suggested by de Waard (1975) to explain the low rates of breast cancer in Japan, and Gray and colleagues (1979) have shown that differences in body size can explain a large portion of the variation in national rates of breast cancer. More recently, Micozzi (1985) has used adult height to serve as an index of energy balance during development and examined the correlation between mean national heights and breast cancer incidence rates (Fig. 14–7); a correlation very similar to that for per capita fat intake was observed.

Within some populations, it does appear legitimate to use height as an index of childhood energy balance. For example, the substantial gain in stature by offspring of Japanese emigrants to the United States provides clear evidence that caloric restriction has occurred in Japan (Insull et al., 1968). As not all members of a society are likely to have been equally restricted, the energy restriction hypothesis would lead us to expect a positive association between height and risk of breast cancer within countries, such as Japan, that have experienced a major secular change in adult height over this century. The positive association between height and breast cancer observed in Greece (Valaoras et al., 1969) and in Holland (de Waard, 1975) may be related to limited energy availability for some girls during periods of social disruption.

The interpretation of height in case-control or cohort studies within countries with a prolonged period of relative affluence is less clear. In the United States, caloric restriction sufficiently severe to limit attained height is likely to be less common so that variability in height largely reflects genetic factors rather than nutritional status during development. Thus, it is not surprising that height has little if any association with risk of breast cancer in most Western industri-

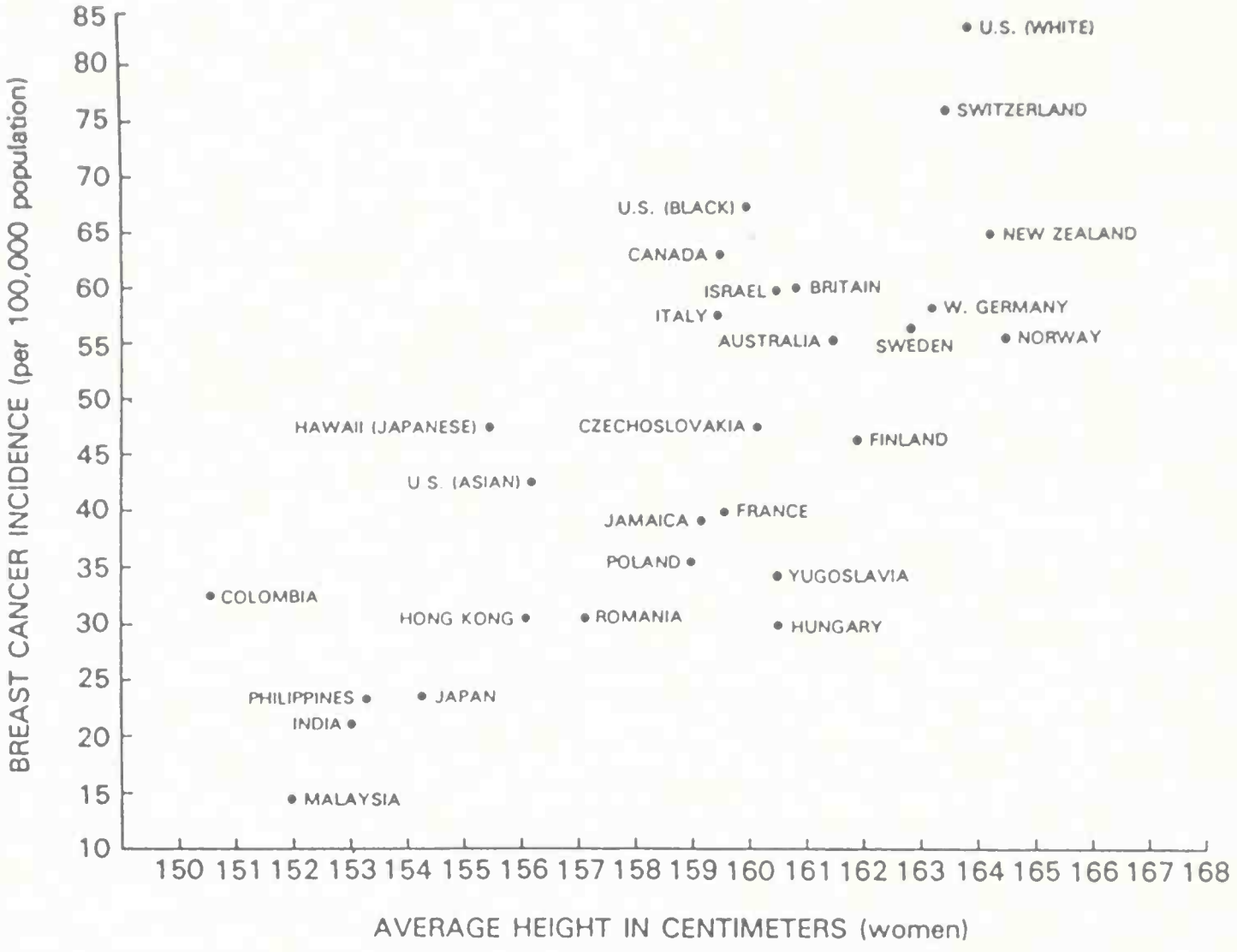


Figure 14–7. Correlation of average adult height in women with breast cancer incidence for 30 countries ($r = 0.8$). (From Micozzi, 1985; reproduced with permission.)

alized countries (Adami et al., 1977; Waaler and Lund, 1983; Willett et al., 1985b).

The relationship between measures of relative weight and breast cancer incidence appears to be particularly complex (Willett, 1987). A full review of this topic is beyond the scope of this chapter; however, among premenopausal women, risk of breast cancer has been inversely related to indices of obesity during adult life and at age 18 (Choi et al., 1978; Paffenbarger et al., 1980; Willett et al., 1985b; Le Marchand et al., 1988). This finding, which is seemingly at odds with the energy-restriction hypothesis, is probably not completely explained by earlier detection of breast tumors among thin women (Willett et al., 1985b) or by an association of leanness with alcohol consumption (Willett et al., 1987b). The observation that the irregularity of menstrual periods is greater among women with a higher relative weight (Willett et al., 1985b) is compatible with the hypothesis of Pike and colleagues (1983) that repeated ovulatory menstrual cycles and the accompanying mitoses of breast tissue increases the likelihood of breast cancer. Among postmenopausal women, a modest positive association has generally been observed between relative weight and risk of breast cancer (Valaoras et al., 1969; Lew and Garfinkel, 1979; Brinton et al., 1979; Helmrich et al., 1983), although a dose-response relationship has not been consistently seen.

The energy-restriction and dietary fat-composition hypotheses are not mutually exclusive and, in the extreme, they may converge. In traditional societies of

physically active farmers who experienced chronic parasitic diseases and recurrent bouts of diarrhea and other infections during childhood, it would be difficult to avoid energy restriction on a rice or other staple diet that contained only 10 percent of calories as fat. Our sedentary life-style, relative control of infectious diseases, and ready availability of refined carbohydrates, however, makes it easy to obtain excess energy on diets that we would consider highly fat-restricted, that is, 20 or 25 percent of calories as fat. Even if it is true that energy restriction sufficiently severe to limit adult height reduces breast cancer incidence, this is unlikely to provide a practical approach to the prevention of human breast cancer. If energy restriction were more moderate, so that only weight but not height were reduced, the apparent inverse relationship of relative weight and breast cancer during the premenopausal years makes in unclear whether the net impact on breast cancer rates would be beneficial, harmful, or neutral. Even if obesity is found to be related to breast cancer rates during the postmenopausal years, the very modest strength of association and relatively limited success of efforts to control obesity make it unlikely that this avenue will lead to a major reduction in this disease.

There are good reasons to reduce intake of animal fat and avoid obesity. Existing data, however, provide little support for the hypothesis that recommended changes in dietary fat composition or a practically attainable restriction of energy intake during adulthood will lead to a substantial reduction in breast cancer in Western cultures.

REFERENCES

Adami, H. O., A. Rimsten, B. Stenkvist, et al. (1977). Influence of height, weight, and obesity on risk of breast cancer in an unselected Swedish population. *Br. J. Cancer 36,* 787–792.

Adelstein, A. M., J. Staszewski, and C. S. Muir (1979). Cancer mortality in 1970–1972 among Polish-born migrants to England and Wales. *Br. J. Cancer 40,* 464–475.

Albanes, D. (1987). Total calories, body weight, and tumor incidence in mice. *Cancer Res. 47,* 1987–1992.

Ames, B. N. (1983). Dietary carcinogens and anticarcinogens. Oxygen radicals and degenerative diseases. *Science 221,* 1256–1264.

Armstrong, B. and R. Doll (1975). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *Int. J. Cancer 15,* 617–631.

Armstrong, B. K., J. B. Brown, H. T. Clarke, et al. (1981). Diet and reproductive hormones: A study of vegetarian and nonvegetarian postmenopausal women. *J.N.C.I. 67,* 761–767.

Beaton, G. H., J. Milner, P. Corey, et al. (1979). Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *Am. J. Clin. Nutr. 32,* 2546–2549.

Birt, D. F. (1986). Dietary fat and experimental carcinogenesis: a summary of recent in vivo studies. In *Advances in Experimental Medicine and Biology: Essential Nutrients in Carcinogenesis,* vol. 206. L. A. Poirer, P. M. Newberne, M. W. Pariza (eds) New York: Plenum Press, pp. 69–84.

Bjarnason, O., N. Day, G. Snaedal, and H. Tulinius (1974). The effect of year of birth on the breast cancer age-incidence curve in Iceland. *Int. J. Cancer 18,* 689–696.

Boissonneault, G. A., C. E. Elson, and M. W. Pariza (1986). Net energy effects of dietary fat on chemically induced mammary carcinogenesis in F344 rats. *J.N.C.I. 76,* 335–338.

Brinton, L. A., R. R. Williams, and R. N. Hoover, et al. (1979). Breast cancer risk factors among screening program participants. *J.N.C.I. 62,* 37–44.

Buell P. (1973). Changing incidence of breast cancer in Japanese-American women. *J.N.C.I. 51,* 1479–1483.

Buring, J. E., C. H. Hennekens, R. J. Lipnick, W. Willett, M. J. Stampfer, B. Rosner, R. Peto, and F. E. Speizer (1987). A prospective cohort study of postmenopausal hormone use and risk of breast cancer in U.S. women. *Am. J. Epidemiol. 125,* 939–947.

Carroll, K. K. (1975). Experimental evidence of dietary factors and hormone-dependent cancers. *Cancer Res. 35,* 3374–3383.

Carroll, K. K. and G. J. Hopkins (1979). Dietary polyunsaturated fat versus saturated fat in relation to mammary carcinogenesis. *Lipids 14,* 155–158.

Chan, P. C. and L. A. Cohen (1974). Effect of dietary fat, antiestrogen, and antiprolactin on the development of mammary tumors in rats. *J.N.C.I. 52,* 25–30.

Choi, N. W., G. R. Howe, A. B. Miller, V. Matthews, R. W. Morgan, L. Munan, J. D. Burch, J. Feather, M. Jain, and A. Kelly (1978). An epidemiologic study of breast cancer. *Am. J. Epidemiol. 107,* 510–521.

Committee on Diet, Nutrition and Cancer (1982). *National Research Council. Diet, Nutrition, and Cancer.* Washington, D.C.: National Academy Press.

de Waard, F., E. A. Baanders-van Halewign, and J. Huizinga (1964). The bimodal age distribution of patients with mammary carcinoma: evidence for the existence of two types of human breast cancer. *Cancer 17,* 141–151.

de Waard, F. (1975). Breast cancer incidence and nutritional status with particular reference to body weight and height. *Cancer Res. 35,* 3351–3356.

Donato, K. A. (1987). Efficiency and utilization of various energy sources for growth. *Am. J. Clin. Nutr. (s) 45,* 168–180.

Enig, M. G., R. J. Munn, and M. Keeney (1978). Dietary fats and cancer trends—a critique. *Fed. Proc. 37,* 2215–2220.

Gaskill, S. P., W. L. McGuire, C. K. Osborne, and M. P. Stern (1979). Breast cancer mortality and diet in the United States. *Cancer Res. 39,* 3628–3637.

Goldin, B. R., H. Adlercreutz, J. T. Dwyer, L. Swenson, J. H. Warram, and S. L. Gorbach (1981). Effect of diet on excretion of estrogens in pre- and postmenopausal women. *Cancer Res. 41,* 3771–3773.

Graham, S. J. Marshall, C. Mettlin, T. Rzepka, T. Nemoto, and T. Byers (1982). Diet in the epidemiology of breast cancer. *Am. J. Epidemiol. 116,* 68–75.

Gray, G. E., M. C. Pike, and B. E. Henderson (1979). Breast cancer incidence and mortality rates in different countries in relation to known risk factors and dietary practices. *Br. J. Cancer 39,* 1–7.

Gray, G. E., P. Williams, V. Gerkins, et al. (1982a). Diet and hormone profiles in teenage girls in four countries at different risk for breast cancer. *Prev. Med. 11,* 103–107.

Gray, G. E., M. C. Pike, T. Hirayama, et al. (1982b). Diet and hormone profiles in teenage girls in four countries at different risk of breast cancer. *Prev. Med. 11,* 108–113.

Greenwald, P., C. Clifford, R. Butrum, and D. C. Iverson (1987). Feasibility studies of a low-fat diet to prevent or retard breast cancer. *Am. J. Clin. Nutr. (S) 45,* 347–353.

Grodin, J. M., P. K. Siiteri, and P. C. MacDonald (1973). Source of estrogen production in postmenopausal women. *J. Clin. Endocrinol. Metab. 36,* 207–214.

Haenszel, W. and M. Kurihara (1968). Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J.N.C.I. 40,* 43–68.

Hagerty, M. A., B. J. Howie, S. Tan, and T. D. Shultz (1987). Effect of low- and high-fat intakes on the hormonal milieu of premenopausal women: A controlled metabolic feeding study (abstr). *Am. J. Clin. Nutr. 47,* 653–659.

Hanai, A. and I. Fujimoto (1982). Cancer incidence in Japan in 1975 and changes in epidemiologic features for cancer in Osaka. *Natl. Cancer Inst. Monogr. 62,* 3–7.

Health and Welfare Statistics Association (Japan). (1987). Indices of Health and Welfare; Trends of National Health 33, No. 9:54, Tokyo.

Helmrich, S. P., S. Shapiro, L. Rosenberg, et al. (1983). Risk factors for breast cancer. *Am. J. Epidemiol. 117,* 35–45.

Hems, G. (1978). The contributions of diet and childbearing to breast cancer rates. *Br. J. Cancer 37,* 974–982.

Hill, P., L. Garbaczewski, P. Helman, J. Huskisson, E. Sporangisa, and E. L. Wynder (1980). Diet, lifestyle, and menstrual activity. *Am. J. Clin. Nutr. 33,* 1192–1198.

Hislop, T. G., A. J. Coldman, J. M. Elwood, G. Brauer, and L. Kan (1986). Childhood and recent eating patterns and risk of breast cancer. *Cancer Detect. Prev. 9,* 47–58.

Hirayama, T. (1978). Epidemiology of breast cancer with special reference to the role of diet. *Prev. Med. 7,* 173–195.

Hirohata, T., T. Shigematsu, A. M. Nomura, Y. Nomura, A. Horie, and I. Hirohata (1985). Occurrence of breast cancer in relation to diet and reproductive history: A case-control study in Fukuoka, Japan. *N.C.I. Monogr. 69,* 187–190.

Hirohata, T., A. M. Nomura, J. H. Hankin, L. N. Kolonel, and J. Lee (1987). An epidemiologic study on the association between diet and breast cancer. *J.N.C.I. 78,* 595–600.

Hopkins, G. J. and K. K. Carroll (1979). Relationship between amount and type of dietary fat in promotion of mammary carcinogenesis induced by 7,12-dimethylbenz(a)anthracene. *J.N.C.I. 62,* 1009–1012.

Howe, G. R. (1985). The use of polytomous dual response data to increase power in case control studies: an application to the association between dietary fat and breast cancer. *J. Chron. Dis. 38,* 663–670.

Ingram, D. M. (1981). Trends in diet and breast cancer mortality in England and Wales 1928–1977. *Nutr. Cancer 3,* 75–80.

Insull, W., T. Oiso, and K. Tsuchiya (1968). Diet and nutritional status of Japanese. *Am. J. Clin. Nutr. 21,* 753–777.

Ip, C. (1987). Fat and essential fatty acid in mammary carcinogensis (Suppl). *Am. J. Clin. Nutr. 45,* 218–224.

Jones, D. Y., A. Schatzkin, S. B. Green, G. Block, L. A. Brinton, R. G. Ziegler, R. Hoover, and P. R. Taylor (1987). Dietary fat and breast cancer in the National Health and Nutrition Examination Survey I. Epidemiologic follow-up study. *J.N.C.I. 79,* 465–471.

Kamiyama, S. and O. Michioka (1983). Mutagenic components of diets in high and low-risk areas for stomach cancer. In Stich HF, ed.: *Carcinogens and Mutagens in the Environment.* Boca Raton, Fla: CRC Press, pp. 29–42.

Karmali, R. A. (1987). Fatty acids: Inhibition (Suppl). *Am. J. Clin. Nutr. 45,* 225–229.

Katsouyanni, K., D. Trichopoulos, V. P. Boyle, E. Xirouchaki, A. Trichopoulou, B. Lisseos, S. Vasilaros, and B. MacMahon (1986). Diet and breast cancer: A case-control study in Greece. *Int. J. Cancer 38,* 815–820.

Katsouyanni, K., W. Willett, D. Trichopoulos, P. Boyle, A. Trichopoulou, S. Vasilaros, J. Papadiamantis, and B. MacMahon (1988). Risk of breast cancer among Greek women in relation to nutrient intake. *Cancer 61,* 181–185.

Key, T. J., S. C. Darby, and M. C. Pike (1987). Trends in breast cancer mortality and diet in England and Wales from 1911 to 1980. *Nutr. Cancer 10,* 1–9.

Kinlen, L. J. (1982). Meat and fat consumption and cancer mortality: A study of strict religious orders in Britain. *Lancet 1,* 946–949.

Knox, E. G. (1977). Foods and diseases. *Br. J. Prev. Soc. Med 31,* 71–80.

Kolonel, L. N., J. H. Hankin, A. M. Nomura, and S. Y. Chu (1981). Dietary fat intake and cancer incidence among five ethnic groups in Hawaii. *Cancer Res. 41,* 3727–3728.

Le, M. G., L. H. Moulton, C. Hill, and A. Kramar (1986). Consumption of dairy produce and alcohol in a case-control study of breast cancer. *J.N.C.I. 77,* 633–636.

Le Marchand, L., L. N. Kolonel, M. E. Earle, and M. P. Mi (1988). Body size at different periods of life and breast cancer risk. *Am. J. Epidemiol. 128,* 137–152.

Lew, E. A. and L. Garfinkel (1979). Variations in mortality by weight among 750,000 men and women. *J. Chronic. Dis. 32,* 563–576.

Longnecker, M. P., J. A. Berlin, M. J. Orza, and T. C. Chalmers (1988). A meta-analysis of alcohol consumption in relation to risk of breast cancer. *J.A.M.A. 260,* 652–656.

Lubin, J. H., P. E. Burns, W. J. Blot, R. G. Ziegler, A. W. Lees, and J. F. Fraumeni, Jr. (1981). Dietary factors and breast cancer risk. *Int. J. Cancer 28,* 685–689.

Lubin, F., Y. Wax, and B. Modan (1986). Role of fat, animal protein, and dietary fiber in breast cancer etiology: A case-control study. *J.N.C.I. 77,* 605–612.

Manousos, O., N. E. Day, D. Trichopoulos, F. Gerovassilis, A. Tzonou, and A. Polychronopoulou (1983). Diet and colorectal cancer: A case-control study in Greece. *Int. J. Cancer 32,* 1–5.

McMichael, A. J. and G. G. Giles (1988). Cancer in migrants to Australia: extending the descriptive epidemiological data. *Cancer Research 48,* 751–756.

Micozzi, M. S. (1985). Nutrition, body size, and breast cancer. *Yearbook Phys. Anthropol. 28,* 175–206.

Miller, A. B., A. Kelly, N. W. Choi, et al. (1978). A study of diet and breast cancer. *Am. J. Epidemiol. 107,* 499–509.

Mills, P. K., J. F. Annegers, and R. L. Phillips (1988). Animal product consumption and subsequent fatal breast cancer among Seventh-Day Adventists. *Am. J. Epidemiol. 127,* 440–453.

Morgan, R. W., M. Jain, A. B. Miller, N. W. Choi, V. Matthews, L. Munan, J. D. Burch, J. Feather, G. R. Howe, and A. Kelly (1978). A comparison of dietary methods in epidemiologic studies. *Am. J. Epidemiol. 107,* 488–498.

National Cancer Institute (1984). Cancer prevention: Good news; better news; best news. Washington, D.C.: Department of Health and Human Services, DHHS publication no. (NIH) 84-2671.

Nukada, A. (1975). Industrialization as a factor for secular increase in physiques of school children in Japan. In Asahina, K. and R. Shigiya (eds.): *Physiological Adaptability and Nutritional Status of the Japanese.* Tokyo: Univ. of Tokyo Press, p. 108.

Paffenbarger, R. S. Jr., J. B. Kampert, and H. C. Chang (1980). Characteristics that predict risk of breast cancer before and after the menopause. *Am. J. Epidemiol. 112,* 258–268.

Pariza, M. W. and R. K. Boutwell. (1987). Historical perspective: Calories and energy expenditure in carcinogenesis (Suppl). *Am. J. Clin. Nutr. 45,* 151–6.

Phillips, R. L. (1975). Role of life-style and dietary habits in risk of cancer among Seventh-Day Adventists. *Cancer Res. 35,* 3513–3522.

Phillips, R. L., L. Garfinkel, J. W. Kuzma, W. L. Beeson, T. Lotz, and B. Brin (1980). Mortality among California Seventh-Day Adventists for selected cancer sites. *J.N.C.I. 65,* 1097–1107.

Phillips, R. L. and D. A. Snowdon (1983). Association of meat and coffee use with cancers of the large bowel, breast, and prostate among Seventh-Day Adventists: Preliminary results. *Cancer Res. 43 (5:Suppl),* 2403s–2408s.

Pike, M. C., M. D. Krailo, B. E. Henderson, J. T. Casagrande, and D. G. Hoel (1983). "Hormonal" risk factors, "breast tissue age" and the age-incidence of breast cancer. *Nature 303,* 767–770.

Potter, J. D. (1987). Dietary fat and the risk of breast cancer (letter). *N. Engl. J. Med. 317,* 166.

Prentice, R. L., F. Kakar, S. Hursting, L. Sheppard, R. Klein, and L. Kushi (1988). Aspects of the rationale for the women's health trial. *J.N.C.I. 80,* 802–814.

Rohan, T. E., and J. D. Potter (1984). Retrospective intake of dietary intake. *Am. J. Epidemiol. 120,* 865–875.

Rohan, T. E., A. J. McMichael, and P. A. Baghurst (1988). A population-based case-control study of diet and breast cancer in Australia. *Am J Epidemiol 128,* 478–489.

Rose, D. P., A. P. Boyar, C. Cohen and L. E. Strong (1987a). Effect of a low-fat diet on hormone levels in women with cystic breast disease. I. Serum steroids and gonadotrophins. *J.N.C.I. 78,* 623–626.

Rose, D. P., L. A. Cohen, B. Berke, and A. P. Boyar (1987b). Effect of a low-fat diet on hormone levels in women with cystic breast disease. II. Serum radioimmunoassayable prolactin and growth hormone and bioactive lactogenic hormones. *J.N.C.I. 78,* 627–631.

Rosner, B. A., W. C. Willett, and D. Speigelman (1989). Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stats in Med* (in press).

Schatzkin, A., Y. Jones, R. N. Hoover, et al. (1987). Alcohol consumption and breast cancer in the epidemiologic follow-up study of the first National Health and Nutrition Examination Survey. *N. Engl. J. Med. 316,* 1169–1173.

Schrauzer, G. N., D. A. White, and C. J. Schneider (1977). Cancer mortality correlation studies. III. Statistical associations with dietary selenium intakes. *Bioinorgan Chem. 7,* 23–31.

Shekelle, R. B., et al. (1987). Re: Total energy intake: Implications for epidemiologic analyses (letter). *Am. J. Epidemiol. 126,* 980.

Sonnenschein, E., L. Glickman, L. McKee, and M. Goldschmidt (1987). Nutritional risk factors for spontaneous breast cancer in pet dogs: a case-control study (abstr). *Am. J. Epidemiol. 126,* 736.

Staszewski, J. and W. Haenszel (1965). Cancer mortality among the Polish-born in the United States. *J.N.C.I. 35,* 291–297.

Stocks, P. (1970). Breast cancer anomalies. *Br. J. Cancer 24,* 633–643.

Talamini, R., C. La Veechia, A. Decarli, et al. (1984). Social factors, diet and breast cancer in a northern Italian population. *Br. J. Cancer 49,* 723–729.

Tannenbaum, A. and H. Silverstone (1953). Nutrition in relation to cancer. *Adv. Cancer Res. 1,* 451–501.

Valaoras, V. G., B. MacMahon, D. Trichopoulos, et al. (1969). Lactation and reproductive histories of breast cancer patients in greater Athens, 1965–67. *Int. J. Cancer 4,* 350–363.

Waaler, H. T. and E. Lund (1983). Association between body height and death from breast cancer. (letter) *Br. J. Cancer 48,* 149–150.

Willett, W. C. (1987). Implications of total energy intake for epidemiological studies of breast and large bowel cancer. *Am J. Clin. Nutr.(s) 45,* 354–360.

Willett, W. C. and B. MacMahon (1984). Diet and cancer—An overview. *N. Engl. J. Med. 310,* 633–638, 697–703.

Willett, W. C., L. Sampson, M. J. Stampfer, et al. (1985a). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol. 122,* 51–65.

Willett, W. C., M. L. Browne, C. Bain, et al. (1985b). Relative weight and risk of breast cancer among premenopausal women. *Am. J. Epidemiol. 122,* 731–740.

Willett, W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, and F. E. Speizer (1987a). Dietary fat and risk of breast cancer. *N. Engl. J. Med. 316,* 22–28.

Willett, W. C., M. J. Stampfer, G. A. Colditz, B. A. Rosner, C. H. Hennekens, and F. E. Speizer (1987). Moderate alcohol consumption and the risk of breast cancer. *N. Engl. J. Med. 316,* 1174–1180.

Willett, W. C., and M. J. Stampfer (1987b). Re: Total energy intake: implications for epidemiologic analyses (letter). *Am. J. Epidemiol. 126,* 982.

Wynder, E. L., F. MacCornack, P. Hill, L. A. Cohen, P. C. Chan, and J. H. Weisburger (1976). Nutrition and the etiology and prevention of breast cancer. *Cancer Detect Prev. 1,* 293–310.

15

Diet and Coronary Heart Disease

Few topics have generated as much scientific controversy during this century as the "diet–heart" hypothesis. According to the classic hypothesis (Fig. 15–1), high intake of saturated fats and cholesterol and low intake of polyunsaturated fats increase the level of serum cholesterol, which leads to development of atheromatous plaques inside arteries. Accumulation of these plaques results in narrowing of the coronary arteries, reduced blood flow to the heart muscle, and finally in the occurrence of myocardial infarction, or "heart attack." The focus of this chapter is to examine the epidemiologic evidence addressing this hypothesis, to discuss the limitations of the hypothesis, and to consider additional hypotheses relating diet to heart disease. Several relevant areas, including vast literatures on animal studies and metabolic experiments relating diet with blood lipids, are mentioned only briefly; these are reviewed in detail elsewhere (Wissler and Vesselinovitch, 1975; Grundy et al., 1982; American Heart Association, 1984; Consensus Conference, 1985).

The inception of the diet–heart hypothesis has been recounted by Gordon (1988); much of the early stimulus derived from the demonstration during the 1930s that dietary cholesterol can cause arterial lesions in animals and that this effect is mediated largely through elevation in plasma cholesterol (Anitschkow, 1967; Katz and Stamler, 1953; Wissler and Vesselinovitch, 1975; Grundy et al., 1982). The further development of the diet–heart hypothesis appears to have been influenced heavily by two lines of epidemiologic evidence, the first being ecologic correlations relating diet to rates of heart disease. These data, along with findings from migrant studies and special populations, are discussed under descriptive studies. The other primary line of evidence derives from studies of serum cholesterol. These studies have related dietary factors to serum cholesterol and, in separate investigations, serum cholesterol to the risk of coronary heart disease. These are discussed briefly in a separate section as such studies do not directly address the relation of diet to heart disease.

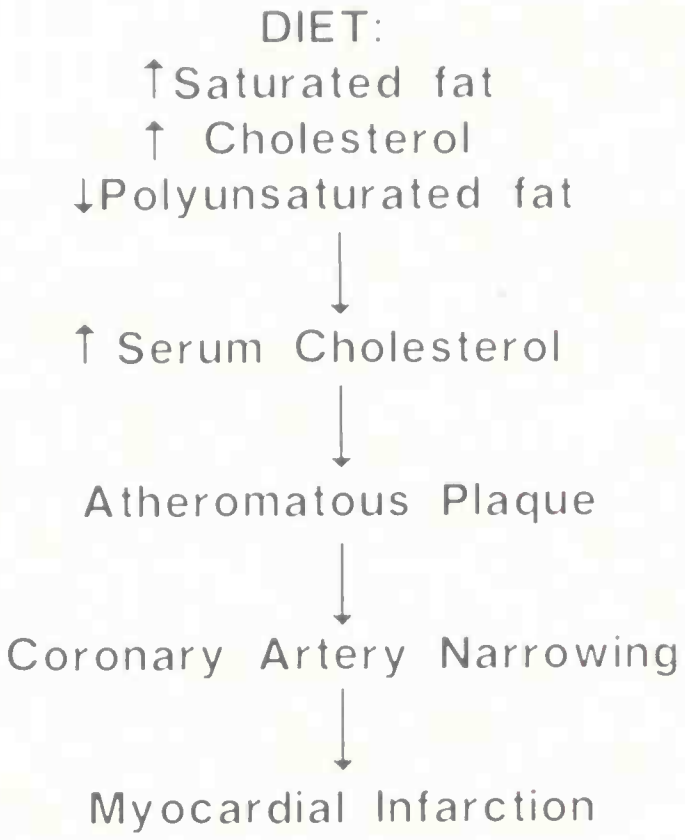


Figure 15–1. Classic diet-heart hypothesis.

DESCRIPTIVE STUDIES OF DIET AND CORONARY HEART DISEASE

The most influential descriptive data on diet and coronary heart disease (CHD) was the work of Keys, relating the mean intake of dietary factors of 16 defined populations in seven countries to the incidence of heart disease in those same groups (Keys, 1980). As shown in Figure 15–2, intake of saturated fat as a percentage of calories was strongly correlated with coronary death rates ($r = 0.84$). In this instance, the countries with low fat intake and low incidence of CHD were the less economically industrialized populations and are likely to have differed in many ways, in particular in physical activity, obesity and, at that time, smoking habits. Indeed, the slope of the line relating saturated fat to risk of CHD death is nearly 2½ times greater than would be expected if the effect of saturated fat operated only by raising serum cholesterol. In another type of ecologic study, a strong association ($r = 0.67$) was observed between the percentage of calories from fat in 12 countries and prevalence of raised atherosclerotic lesions in autopsy cases from the same geographic area (Scrimshaw and Guzman, 1968). Little correlation ($r = 0.07$), however, was seen with the percentage of total fat from animal fat. As with any correlational study, the possibility of confounding by other risk factors exists. The clearest message from these data and other comparisons (McGill, 1968) is that rates of CHD differ dramatically among countries and that the highly developed countries are at highest risk. Other correlational studies are cited in the review by Grundy and co-workers (1982).

Migrant Studies

Coronary heart disease incidence rates among three defined Japanese populations living in Japan, Hawaii, and San Francisco have been compared (Kato et

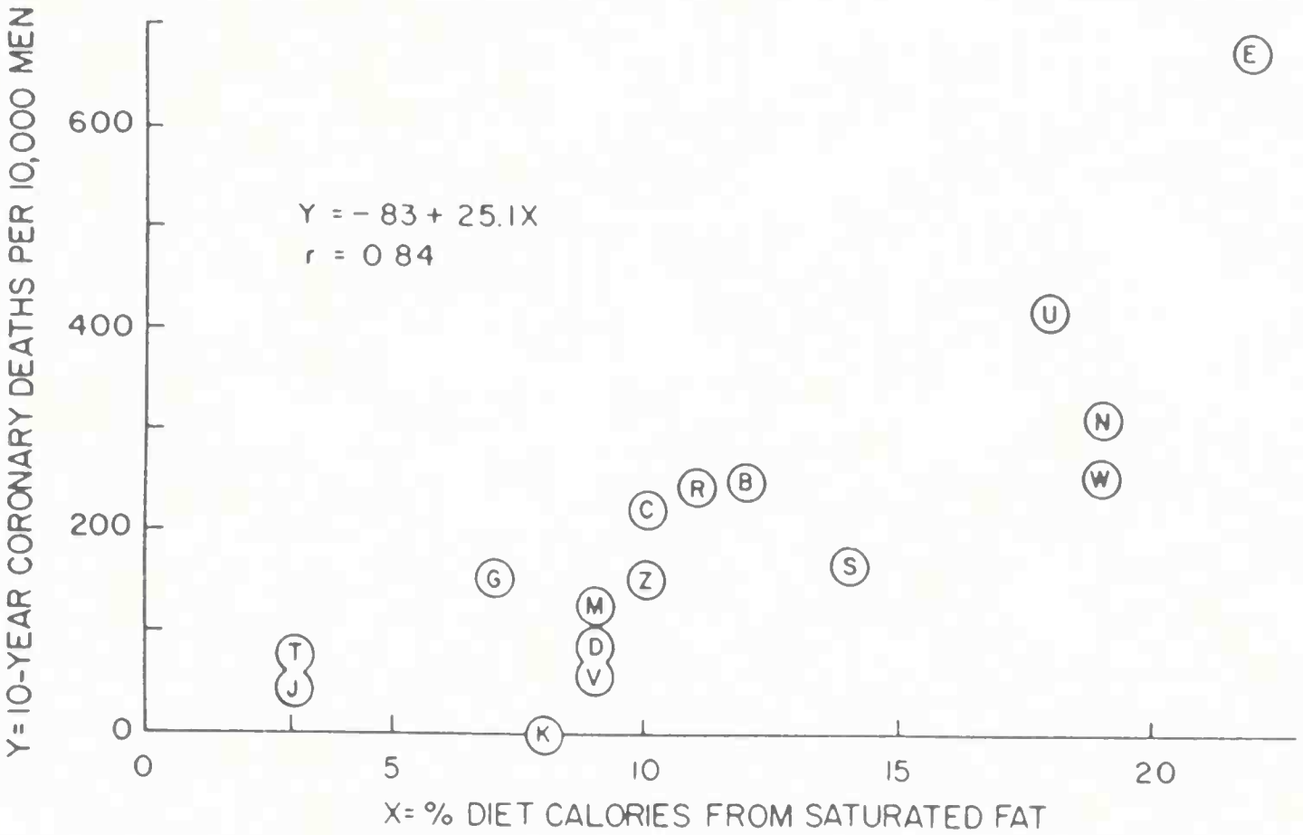


Figure 15–2. Ten-year coronary death rates of the cohorts plotted against the percentage of dietary calories supplied by saturated fatty acids. (From Keys, 1980; reproduced with permission.)

al., 1973; Robertson et al., 1977). Saturated fat intake as a percentage of calories in the three populations was found to be 7, 23, and 26 percent, respectively. For these three groups, mean relative weight (but not height) and serum cholesterol increased in parallel with the saturated fat intake, mean alcohol intake markedly decreased, and dietary cholesterol was similar. Age-adjusted CHD incidence rates were 1.6 per 1000 person-years in Japan, 3.0 in Hawaii, and 3.7 in San Francisco. These data indicate that the substantial differences in rates of CHD among the three areas cannot be explained by genetic factors and are consistent with the hypothesis that dietary saturated fat may contribute to the difference. Because of other dietary and nondietary factors, such as alcohol intake and obesity, that vary among the groups, specific causal factors cannot be firmly identified on the basis of this type of data alone.

Secular Trends

Data on secular trends in CHD mortality are of interest, but do not provide clear support for the diet–heart hypothesis. Within North America, CHD mortality rates rose dramatically over the first half of this century (Anderson, 1979), although the increase in total and saturated fat consumption was slight and polyunsaturated fat rose two- to threefold (Friend, 1967; Page and Marston, 1979). Kahn (1970) has estimated that these changes in food consumption would account for only a 10-mg/dl increase in serum cholesterol and an 8 percent rise in incidence of CHD. Since the late 1960s age-specific mortality rates have decreased steadily by a total of approximately 30 percent for persons aged 35 to 74 years (Working Group, 1981; Stamler, 1985). During the period, dairy fat and

lard have been partially replaced by vegetable fat with the overall effect of increasing polyunsaturated fat intake and slightly decreasing saturated fat intake (Page and Marston, 1979; Stamler, 1985). During World War II, major reductions in total and animal fat consumption occurred in Finland, Norway, and other areas in Europe. Clear decreases in deaths ascribed to arteriosclerosis occurred during this period, similar to decreases observed during World War I (Katz and Stamler, 1953).

In Japan, saturated fat and cholesterol intake has increased dramatically since 1950 and the change in CHD mortality has been large, approximately 2½-fold between 1950 and 1968 (Wen and Gershoff, 1973). Wen and Gershoff, however, estimated that less than one-fourth of the increase in CHD mortality could be accounted for by dietary changes in saturated and unsaturated fats and cholesterol if it is assumed that their effect is mediated only by serum cholesterol.

Eskin (1971) observed that per capita egg consumption tripled during the period 1920 to 1970 in the United Kingdom, whereas CHD death rates rose approximately fivefold during this period, an increase that would far exceed the effect explainable by the predicted influence of the cholesterol content of eggs on serum cholesterol.

In summary, the studies of diet and coronary heart disease that use population groups as units of observation provide convincing evidence that rates of disease differ dramatically among populations and that these differences are primarily due to environmental rather than genetic factors. Although these studies tend to support the classic diet–heart hypothesis, alternative explanations cannot be excluded with confidence.

STUDIES OF BLOOD CHOLESTEROL AS AN INTERMEDIARY FACTOR

Studies of blood cholesterol as a factor that may mediate the relationship between diet and CHD are considered in two parts: those relating diet to blood cholesterol and those relating blood cholesterol to CHD.

Studies of Diet and Blood Cholesterol

Relationships between dietary factors and blood lipid levels are best studied by controlled, preferably blinded, feeding trials in human subjects. The number of subjects required is usually small (typically 10 to 50), and hundreds of such studies have been conducted. Although there continues to be some debate about details of the shape of the dose-response relationships (Keys, 1984; Hegsted, 1986), there is no question that higher intake of dietary cholesterol and saturated fats and lower intake of polyunsaturated fats increase blood total cholesterol levels. Data from many metabolic studies have been summarized by Hegsted and colleagues (1965) and Keys and Parlin (1966) as equations that predict serum total cholesterol from intake of saturated fat, polyunsaturated fat, and cholesterol (see Chapter 9 for these equations). Despite the clearly described quanti-

tative relationships between these dietary factors and serum cholesterol, the specific mechanisms by which they influence blood cholesterol has not been firmly established.

The relation of dietary lipids to blood cholesterol has also been studied cross-sectionally in human populations. For example, intake of fat-containing foods assessed using a simple frequency questionnaire was not associated with serum cholesterol among a population of men participating in the Tecumseh Heart Study (Nichols et al., 1976); this has sometimes been cited as evidence against the diet–heart hypothesis (Mann, 1977). Except for exploratory types of analyses, however, a cross-sectional study design is not optimal to define the effects of dietary factors on blood lipids. As noted in Chapter 1 and the work of Jacobs and colleagues (1979), the true correlation of dietary fat intake with a single blood measurement of total cholesterol within a general U.S. population is expected to be very low because calculations based on the results of controlled metabolic studies indicate that most of the between-person variation in cholesterol is unrelated to dietary fats. These calculations, coupled with some inevitable degree of measurement error, indicate that a very large sample size would be needed to detect any positive correlation between diet and blood cholesterol. An association has been seen between dietary fat intake, expressed as the Keys score, and serum total cholesterol among the nearly 2000 men participating in the Western Electric Study (Shekelle et al., 1981), but this correlation was, as expected, low ($r = 0.08$). The Keys equation indicates that the relation between cholesterol intake and blood level is nonlinear, being related to the square root of cholesterol intake; thus, it might be expected that a stronger correlation would be seen among populations with a lower cholesterol distribution. This seems to be the case as a relatively strong correlation has been observed between cholesterol intake and blood level among vegetarians (Kushi et al., 1988) and an Indian population living in Mexico (Conner et al., 1978); both groups consumed relatively low amounts of cholesterol.

Blood Cholesterol Level and Risk of Coronary Heart Disease

Few relationships are as well established as the association between blood total cholesterol levels and risk of CHD. The number of studies in which this has been observed are too numerous to review or enumerate here. The findings from the largest of these studies, based on analyses from 356,222 men screened for participation in the Multiple Risk Factor Intervention Trial (MRFIT) (1982) study, are displayed in Table 15–1 (Stamler et al., 1986). As can be seen, the relationship is continuous and without threshold or cut-off over a wide range of cholesterol values.

Because the relationship between serum cholesterol and incidence of coronary heart disease found in MRFIT and most other studies was based on a single cholesterol determination, an appreciable degree of measurement error will have occurred. The effect of this misclassification is to attenuate or reduce the relationship between true long-term average serum cholesterol for an individual and risk of disease. The magnitude of this attenuation can be estimated from the

Table 15–1. Relation between serum cholesterol and 6-year coronary heart disease mortality among 356,222 screenees in MRFIT

Decile	Serum cholesterol, mg/dl (mmol/L)	Mean serum cholesterol, mg/dl (mmol/L)	CHD mortality		
			No. of deaths	Rate per 1000	Relative risk
1	≤167 (≤4.32)	153.2 (3.962)	95	3.16	1.00
2	168–181 (4.34–4.68)	175.0 (4.526)	101	3.32	1.05
3	182–192 (4.71–4.97)	187.1 (4.838)	139	4.15	1.31
4	193–202 (4.99–5.22)	197.6 (5.110)	149	4.21	1.33
5	203–212 (5.25–5.48)	207.5 (5.366)	203	5.43	1.72
6	213–220 (5.51–5.69)	216.1 (5.588)	192	5.81	1.84
7	221–231 (5.72–5.97)	225.9 (5.842)	261	6.94	2.20
8	232–244 (6.00–6.31)	237.7 (6.147)	272	7.35	2.33
9	245–263 (6.34–6.80)	253.4 (6.553)	352	9.10	2.88
10	≥264 (≥6.83)	289.5 (7.486)	494	13.05	4.13

CHD indicates coronary heart disease; MRFIT = Multiple Risk Factor Intervention Trial. Analysis is age standardized.

Stamler et al., 1986; reproduced with permission from the publisher.

intraclass correlation for repeated serum cholesterol measurements, which is equivalent to the regression coefficient for a single measurement (independent variable) predicting the true long-term average level (dependent variable). In the study of Shekelle and co-workers (1981), this coefficient (γ) for repeated measures at a 2-year interval was 0.65. As described in Chapter 12, γ can be used to approximate the true relative risk unattenuated by measurement error, specifically, true relative risk = $\exp(\beta/\gamma)$ where β is the observed logistic regression coefficient. From the study of Stamler and co-workers (1986), the observed relative risk for a 20-mg/dl difference in serum cholesterol was 1.17. Thus, the true relative risk for a long-term difference of 20 mg/dl in serum cholesterol would be approximately $\exp(\ln 1.17/0.65) = 1.27$. One implication of the attenuated association observed in typical studies is that the attempts described in the section on secular trends to predict effects of a dietary change on CHD rates have underestimated the potential impact of diet because the assumed relationship between serum cholesterol and CHD was erroneously weak.

The consistency of a clear dose-response relationship seen in a multitude of studies and the existence of a plausible mechanism leaves little doubt that the association between blood cholesterol and risk of CHD is causal. This conclusion was strengthened by the findings of the Lipid Research Clinics trial in which men with high blood cholesterol levels randomized to cholestyramine, a lipid lowering agent, experienced a marginally significant reduction in incidence of myocardial infarction compared with a placebo group (Lipid Research Clinics Program, 1984). An even greater reduction in incidence was observed in a trial of gemfibrozil, another lipid lowering drug, that was conducted among a population of men with only moderately elevated blood cholesterol levels (Frick et al., 1987). In addition to substantiating the causal association between blood cholesterol levels and CHD, these studies documented that the excess risk can be reversed, at least in part, within relatively few years.

CASE-CONTROL AND COHORT STUDIES OF DIET AND CORONARY HEART DISEASE

The soundly documented relationships between dietary lipids and blood cholesterol and between blood cholesterol and CHD strongly suggest that dietary lipids influence the risk of CHD. Nevertheless, this conclusion cannot be drawn with certainty from such data due to the complex and incompletely understood pathophysiology of CHD. For example, the manipulation of a dietary factor may decrease blood total cholesterol, and thus appear to be beneficial, but have a counterbalancing deleterious effect on another physiologic parameter (discussed later). In other words, although the reduction of blood cholesterol is in general desirable, it is possible that not all methods of reduction produce the same effect on heart disease. Even if a change in dietary lipids influences the incidence of CHD in the direction predicted by its effect on blood cholesterol, the quantitative relationship between this dietary change and risk of disease is uncertain because of the possibility of many other potential physiologic effects of this dietary manipulation. The diet–heart hypothesis would thus be strengthened by demonstrating that individuals who consume more saturated fat and cholesterol and less polyunsaturated fat actually have a higher risk of CHD.

The strongest evidence for the diet–heart hypothesis would be derived from randomized, ideally double-blinded, trials of dietary manipulation in a human population initially free of CHD. In the customary development of a hypothesis, however, such a trial would be preceded by case-control or cohort studies that, if an association with diet were seen, would provide further justification for the expense of a long-term trial. For the diet–heart disease hypothesis, however, trials were pilot-tested or initiated without evidence from case-control or cohort studies. This leap to trials may relate to the strength of the belief in the hypothesis as well as to the failure to find an association between dietary lipids and serum cholesterol in many cross-sectional studies. This lack of association, as noted earlier, has been attributed to homogeneity of diet among the general population and an inability to measure the diets of free-living subjects, rather than to the relative unresponsiveness of serum cholesterol to moderate changes in diet. Whatever the reasons, the fact is that relatively few case-control or cohort studies of diet and CHD have been published and most of these are quite recent. Furthermore, many of these were not specifically designed to be studies of diet and heart disease, but rather are based on data collected for different uses in the past.

Case-Control Studies of Diet and Coronary Heart Disease

Only two case-control studies of diet and CHD seem to have been published, both during the 1960s, and are noted primarily for historical interest. Meredith and colleagues (1960) interviewed 162 North Dakota men with incident or recurring CHD and compared their prehospitalization diets with those of 324 age-matched population controls. Dietary intake was collected using a 135-item

food-frequency questionnaire that included information on portion sizes and methods of preparation. Cases reported higher caloric intake and lower percentages of calories from animal and total fat than controls, but these differences were slight and not statistically significant. No appreciable differences were seen for other dietary lipids and other aspects of the diet were not reported.

Finegan and colleagues (1968) interviewed 100 Irish men hospitalized for CHD using a dietary history interview focused on a hypothetical "average week" and compared responses with those from 50 men admitted for minor surgical problems. Reported caloric intake was nearly identical in the two groups and animal fat as a percentage of calories was slightly higher for cases (36%) compared with controls (33%), but this difference was apparently not statistically significant. As was customary for epidemiologic studies conducted at that time, data were only presented as mean values for cases and controls rather than as relative risks; however, no appreicable differences were seen for polyunsaturated fat, protein, and cholesterol.

In many respects, these were well done case-control studies and the dietary data appear to have been rather comprehensive and carefully collected even though no information on their validity was obtained. Although no significant associations were seen, these studies are far from definitively null as they were rather small and confidence limits would have been wide had they been calculated. No information was provided on the true variation in dietary intake within the study populations, and the possibility of distortions due to biased recall of diet are almost impossible to exclude in the context of a case-control study, especially among a recently hospitalized group.

Prospective Cohort Studies of Diet and Coronary Heart Disease

Because issues of biases related to selection of control subjects and the recall of past diet are eliminated in prospective cohort studies, these investigations should provide more consistent findings on diet and CHD. Most of the available studies, however, were not primarily designed as investigations of diet and heart disease and thus have many limitations. These studies are summarized in Table 15–2 and are briefly discussed here.

Between 1956 and 1966, Morris and colleagues (1977) collected information on food intake using 1-week weighed dietary records as part of an effort to develop a simple questionnaire that could be used in epidemiologic studies. Their approach, using a simple predictor score based on relatively few foods (Heady, 1961), was felt to be unpromising and the effort was abandoned. Nevertheless, they continued to follow the participants for up to 20 years and later analyzed the incidence of fatal or nonfatal myocardial infarction, based on 45 cases, in relation to their earlier diet. A strong inverse association was found between total energy intake and risk of CHD; the relative risk was 0.3 for those in the highest third compared with those in the lowest third. This relationship was particularly strong during the first 5 years of follow-up, for which the relative risk was 0.1. An inverse association was also seen with intake of dietary fiber; when fiber intake was further subdivided according to its food source, the inverse relationship was found to be entirely attributable to cereals rather than

to fiber from fruits, vegetables, legumes, or nuts. No association was observed with percentage of calories from saturated fat or with cholesterol intake divided by calories; however, a nonsignificant inverse trend was noted with the ratio of polyunsaturated to saturated fatty acids, which was particularly strong during the first 5 years of follow-up.

Shekelle and colleagues (1981) used the average of two detailed diet history interviews, each of which included information on the use of 195 foods consumed during the previous 28 days, as baseline information for the follow-up of 1900 men employed by the Western Electric Company. During the next 20 years approximately 200 men (the exact number is not given) died of CHD. Dietary intakes of saturated fat, polyunsaturated fat, and cholesterol were combined as scores for predicting serum cholesterol as described by Keys and by Hegsted (see Chapter 9); a highly significant positive association was observed with these scores. Notably, this effect was seen even with serum cholesterol included in a multivariate model, indicating an influence that was not mediated by serum cholesterol. From the data provided, it can be estimated that the relative risk was 1.25 for a one standard deviation change in the Keys score. Although this finding provides general support for the diet–heart hypothesis, practical decisions regarding diet must be based on information about the individual components of the predictor scores. In this study, specific nutrients were evaluated as nutrient densities. The overall positive association with the predictor score was primarily the result of an inverse relationship with polyunsaturated fat intake and a positive association with dietary cholesterol; saturated fat was not significantly related with CHD risk. Because of the detailed dietary data collected in this investigation and the relatively large population, this is the most informative single study of diet and CHD to the present.

In 1965, a single 24-hour recall was incorporated in three populations (Framingham, Massachusetts; Puerto Rico; and Honolulu, Hawaii) being followed prospectively to identify determinants of CHD. Follow-up analyses have been presented for each center individually (Garcia-Palmieri et al., 1980; Dawber et al., 1982; McGee et al., 1984) and also analyzed jointly by Gordon and co-workers (1981). An inverse relation between total energy intake and incidence of CHD was found in all three studies. An inverse association with starch intake was also found in the Puerto Rican study and positive associations were seen in the Hawaiian data with saturated fat and cholesterol intakes when presented as nutrient densities. Inverse associations with alcohol intake were observed in Framingham and Hawaii and a similar trend, although not quite statistically significant, was seen in Puerto Rico. In a further analysis of the Framingham data, Dawber and colleagues (1982) found no association between egg consumption and incidence of CHD despite a tenfold difference in average consumption between the first and third tertiles.

In 1960, detailed information on usual diet during the previous 6 to 12 months was collected by interview of both participants and their spouses for 857 men free of CHD living in Zutphen, the Netherlands, one of the centers of the Seven-Countries study organized by Keys (Kromhout et al., 1985). During the next 10 years, 30 participants died of CHD. For men in the highest quintile of total energy intake the relative risk of CHD was only about 0.2 compared with

Table 15–2. Prospective cohort studies of dietary factors in relation to risk of coronary heart disease

Study	Population	Dietary method	CHD cases	Energy	Saturated fat	Polyunsaturated fat	Cholesterol	Lipid[a] score	Fiber	Fish	Alcohol	Comments
Morris et al. (1977)	337 U.K. bank clerks	7-day record, weighted	45	↓	0	—	0	—	↓	—	—	Trend of ↓ risk with high P/S ratio
Shekelle et al. (1981 and 1985)	1900 U.S. men	Diet history interview	~200 CHD deaths	—	0	↓	↑	↑	—	↓	—	
Garcia-Palmieri et al. (1980)	8218 Puerto Rican men	24-hour recall	163	↓	0	0	0	—	—	—	↓	Inverse relation with starch intake
Gordon et al. (1981), Dawber et al. (1982)	895 Framingham men	24-hour recall	51	↓	0	0	0	—	—	—	↓	No association with egg intake
McGee et al. (1984)	7088 Honolulu men	24-hour recall	309	↓	↑	0	↑	—	—	—	↓	Dietary fat values were divided by calories
Kromhout et al. (1982, 1984, 1985)	857 Zutphen men (Dutch)	Diet history interview	30 CHD deaths	↓	0	0	0	—	↓	↓	0	Inverse relation with fiber not significant when divided by calories

Kushi et al. (1985)	1001 Irish and Boston men	Diet history interview	110 CHD deaths	0	↑	0	0	↑	↓	—	—	Dietary fat values were divided by calories
Snowdon et al. (1984)	25,153 U.S. Seventh Day Adventists	28-item frequency questionnaire	1599	—	—	—	—	—	—	—	—	Positive association with meat intake (RR = 1.5)
Khaw et al. (1987)	California men and women	24-hour recall	65 CHD deaths	0	—	—	0	—	↓	—	0	
Burr et al. (1982)	10,943 Welch vegetarians	short food-frequency questionnaire	585		—	—	—	—	0	—	—	Vegetarians have lower CHD mortality
Lapidus et al. (1986)	1462 Swedish women	24-hour recall	28 infarctions	↓	—	—	—	—	—	—	—	
Norell et al. (1986)	10,966 Swedish men and women	?	800 CHD deaths	—	—	—	—	—	—	↓	—	

[a]Lipid score refers to Keys or Hegsted scores for predicting serum cholesterol.

↓ = inverse association; ↑ = positive association; 0 = no statistically significant association; — = no information

the those in the lowest quintile. Within this same study population, an inverse relationship was also seen with fiber intake (Kromhout et al., 1984), although this did not remain significant when expressed as a ratio with calories. In addition, men consuming greater quantities of fish experienced a lower rate of death due to CHD (Kromhout et al., 1985).

Kushi and colleagues (1985) traced the mortality experience of 1001 men living in Ireland and Boston who had completed a dietary history interview approximately 20 years earlier. During this period 110 men died of CHD. An inverse association between total energy intake and risk of CHD was observed; however, this did not attain statistical significance. Positive associations were seen with the predictor scores of Keys or Hegsted, which appeared to be largely attributable to a positive association with saturated fat intake. A marginally significant inverse association between fiber intake and risk of CHD death was also seen when expressed as a nutrient density.

In 1968, 1462 Swedish women provided a 24-hour recall as part of prospective study of risk factors for cardiovascular disease (Lapidus et al., 1986). During 12 years of follow-up, 28 women developed either fatal or nonfatal myocardial infarction. A very strong inverse relationship was found between total energy intake and risk of fatal or nonfatal myocardial infarction; this persisted in multivariate analysis including obesity and classical risk factors. No association, however, was observed for intakes of protein, fat, and carbohydrate expressed as a percent of calories.

Khaw and Barrett-Connor (1987) followed 859 men and women who completed a single 24-hour recall in 1972 to 1974 for 12 years during which time 65 persons died of CHD. After adjusting for fiber intake, which was inversely related to risk of CHD, dietary lipids were not significantly associated with outcome.

In addition to the studies described previously in which an assessment of diet was obtained that allowed the calculation of nutrients, several other studies have been reported in which only a limited number of foods were examined. Snowdon and colleagues (1984) examined the association between meat intake, assessed using a short food-frequency questionnaire, and 1599 subsequent deaths due to CHD among a population of 25,153 California Seventh Day Adventists. The risk of CHD death was 1.5 times greater for nonvegetarians compared with vegetarians, and within the nonvegetarian group a dose-response relationship was seen with frequency of meat intake among men but not among women. Burr and Sweetman (1982) followed a cohort of nearly 11,000 men and women who were identified by their interest in "health-foods" and who completed a short but poorly described dietary questionnaire. Although vegetarians had a lower CHD mortality rate than nonvegetarians, neither fish intake or use of wholemeal bread was related to this outcome. Norell and colleagues (1986) collected information regarding fish intake by an unspecified method among 10,966 Swedish men and women. During 14 years of follow-up an inverse relationship with myocardial infarction was found; for those comsuming "high" amounts of fish, the relative risk was 0.70 compared with those consuming "low" amounts.

Summary of Prospective Cohort Studies of Diet and Coronary Heart Disease

Viewed together, the only consistent finding in the prospective studies of diet and CHD is that men or women with higher total energy intake experience lower rates of disease. The interpretation of this finding, as discussed in Chapter 11, is not that an overall increase in food consumption will lead to a reduced risk of CHD, particularly as obesity is positively related to rates of this disease. As pointed out by Morris and colleagues (1977), this inverse relationship almost surely represents a protective effect of physical activity; increased physical activity leads to a higher energy intake and reduced risk of CHD. Thus, it is ironic that the existing studies of diet and CHD provide better support for the influence of exercise than they do for any primary effect of diet. Although the reported studies do provide some general support for the classic diet–heart hypothesis, data for specific dietary lipids is weak and inconsistent; a positive association with saturated fat intake was seen in only two studies and not in the study of Shekelle and colleagues (1981). A positive association with cholesterol intake was found in only two studies, and an inverse relationship with polyunsaturated fat intake in only one.

In addition to these inconsistencies, the interpretation of the published findings is made difficult by the strong inverse association of total energy intake with risk of CHD observed in most of the studies. As pointed out in Chapter 11, this relationship causes specific nutrients to be inversely associated with risk of disease, even if the composition of the diet has no effect. Many authors have recognized this problem and attempted to address it by presenting their findings as nutrient densities, that is, nutrient intake divided by total energy. Dividing by a confounding variable, such as caloric intake, however, does not control confounding; instead this tends to make variables positively associated with risk of CHD (see Table 11–4). Unfortunately, none of the original publications used a theoretically more appropriate analysis. Shekelle and colleagues (1987) have commented that they, in principle, also favor the use of regression analysis but that the use of nutrient densities may not result in substantial error in every instance, noting that the results of their study (Shekelle et al., 1981) were essentially unaltered when using regression analysis. Nonetheless, as they note, this cannot necessarily be assumed to be true for all studies; the distortion depends on the strength of the association between energy intake and disease and between the specific nutrient and energy intake. In a reanalysis of his original data using residuals from regression analyses, Kushi (1987) found that the methods of adjustment for total energy intake did influence his findings; although the association with saturated fat was unchanged, the use of residuals weakened the positive association with cholesterol intake ($p = 0.11$ compared with $p = 0.02$ for nutrient density) and the inverse association with fiber intake ($p = 0.17$ compared with $p = 0.08$ for nutrient density). In the study of Kushi and co-workers, these differences occurred despite a relatively weak association between total energy intake and CHD. Thus, although some published findings based on nutrient densities may not be materially distorted, there is no way of knowing until

appropriate analyses are performed. Although consistent findings for dietary lipids have not been found in the prospective investigations published thus far, relationships with other aspects of diet have been seen more frequently and are discussed later.

Existing studies do not provide consistent findings for any specific dietary lipid, however, they should not be interpreted as providing strong negative evidence. Many of these studies had major limitations, largely stemming from the fact that most were opportunistic analyses of data not originally intended to be used for that purpose. The most obvious limitation of the published studies is their small size; five of these had fewer than 100 endpoints. This limitation is particularly serious as associations are likely to be of modest magnitude (discussed later). The effect of small study size can be readily appreciated when confidence intervals are displayed as potentially important associations are usually included within the interval; however, in many reports these intervals were not provided.

A second limitation of many studies relates to the method of dietary assessment. In a number of studies a single 24-hour recall was used to represent exposure for a period of up to 12 years. Although this does provide some information on diet, as indicated by the consistent finding of an inverse relationship with total energy intake, the degree of misclassification can be substantial and varies among nutrients (see Chapter 3).

A third limitation of many studies is the prolonged follow-up period. Although a long follow-up period is frequently viewed as an advantage, this would not be true if dietary exposure tended to change over time and recent exposure is relevant to risk of disease; this shorter time frame was suggested by an effect of cholesterol reduction within several years in the Lipid Research Clinics cholestyramine trial (Lipid Research Clinics Program, 1984). With prolonged follow-up, misclassification of dietary exposures increases over time. Furthermore, this problem is exacerbated by the tendency for most of the disease in a cohort to occur toward the later part of the follow-up period as the effects of selecting a healthy population at entry will have waned and the participants will have aged. The most desirable solution is to assess diet repeatedly during the follow-up period; alternate analyses can then be conducted with variable lag periods between measurement of dietary intake and occurrence of CHD and between degrees of change in diet and CHD. Unfortunately, none of the studies reported to date have included reassessments of dietary intake. Another approach if one measurement is available, is to subdivide the follow-up period in the analysis of the data; however, this exacerbates the problem of small study size. Such an analytic approach was used in only one of the published studies; Morris and co-workers (1977) found that associations with diet generally tended to be stronger for disease occurring during the first 5 years than for disease occurring later.

The possibility of detecting an association between dietary factors and risk of CHD in a cohort study depends strongly on the degree of variation in dietary intake among study participants; a perceived lack of heterogeneity within the U.S. population has led some to believe that associations will not be found. Many lines of evidence, however, indicate that dietary heterogeneity does exist

within U.S. populations, including the findings of reasonable correlations between nutrient intake assessed by independent methods (see Chapter 6) and the associations, even if not consistent for some nutrients, between specific dietary factors and risk of CHD in the prospective studies cited earlier. The degree of variation, however, is modest for many nutrients, with the result that associations with disease rates are also likely to be modest. Potential magnitudes of association can be estimated from the data provided in Table 1 of the study by Shekelle and colleagues (1981): the standard deviation for the Keys score (which represents the expected variation in serum cholesterol that can be attributed to dietary saturated and polyunsaturated fat and cholesterol) is 8.3 mg/dl. The difference between the 5th and the 95th percentile would be about 3.3 standard deviations or a 27 mg/dl difference in serum cholesterol attributable to these three dietary factors in combination. The degree to which such a difference in serum cholesterol can be translated into risk of CHD can be estimated from the experience of the Lipid Research Clinics trial (Lipid Research Clinics Program, 1984). In that study a 20-mg/dl difference in serum cholesterol corresponded to a 19 percent reduction in risk of CHD incidence. Thus a difference of 27 mg/dl in serum cholesterol might be expected to translate to a difference in risk of approximately 25 percent, meaning that the expected relative risk might be on the order of 1.3. This estimate involves several oversimplifications, including the assumptions that the full effect of intervention was present during the early years of the trial and is entirely mediated by total serum cholesterol. These assumptions, which tend to underestimate the anticipated effect of diet, of course, are counterbalanced by some inevitable degree of misclassification in the measurement of dietary intake. Interestingly, the relative risk observed in the study by Shekelle and colleagues was 1.4 for the highest tertile of Keys score compared with the lowest, a finding somewhat larger than that predicted. Most of the other prospective studies that failed to find an association between dietary lipids and risk of CHD were too small to have excluded a relative risk of this magnitude. Clearly, studies to quantify the relationships of diet with CHD realistically will need to be large.

STUDIES USING BIOCHEMICAL MARKERS OF FATTY ACIDS INTAKE AND RISK OF CORONARY HEART DISEASE

As an alternative to measuring dietary intake, some investigators have used biochemical analyses of plasma lipid fractions, platelet or red cell membranes, or subcutaneous fat as markers of fatty acid intake (see Chapter 9) to study relationships with risk of CHD (see Wood et al., 1987 for a review). In an ecologic study comparing regional mortality rates from CHD and mean levels of fatty acids in adipose of persons from these areas, Riemersma and colleagues (1986) found the lowest levels of adipose linoleic acid in North Karelia, Finland, where CHD mortality was highest, and the highest levels in Italy, where CHD mortality was lowest. Lower levels of adipose linoleic acid were found in 75 men with myocardial infarction compared with 25 controls (Kirkeby et al., 1972a), but

levels were similar in British men dying of CHD and those dying of other causes (Thomas and Scott, 1981). In a cross-sectional study of Scottish men, Wood and colleagues (1984) found significantly lower levels of linoleic acid in the 28 men with previously unidentified CHD than in the healthy men who provided dietary records. In a case-control study including 80 incident cases of myocardial infarction, this same group found that the risk of infarction decreased with higher levels of adipose linoleic acid (Wood et al., 1987); however, cigarette smokers had lower levels of linoleic acid than nonsmokers and adjustment for smoking in a multivariate analysis eliminated the association between linoleic acid and myocardial infarction. In a group of healthy men who provided dietary records, these authors found that smokers consumed a substantially lower percentage of their fat as linoleic acid compared with nonsmokers, thus indicating that lower adipose levels in smokers were not simply a metabolic effect of smoking. Risk of myocardial infarction also increased with higher adipose levels of monounsaturated fats, including both oleic and palmitoleic acids. In the same study, patients with angina pectoris had lower adipose linoleic acid than healthy controls; this inverse relationship was reduced, but remained statistically significant after adjusting for smoking. The rather low participation rate for control subjects, approximately 50 percent, limits the interpretation of findings of this study.

In one small case-control study, platelet membrane linoleic acid was found to be lower in subjects with myocardial infarction compared with controls (Renaud et al., 1970). Lower levels of linoleic acid in plasma or serum cholesteryl esters have been observed in patients with CHD (Lewis, 1958; Schrade et al., 1961; Kirkeby, 1972b). In one case-control study, a lower red-cell membrane level of linoleic acid was found in patients with myocardial infarction (Simpson et al., 1982), although this was not confirmed in another investigation (Lea et al., 1982).

In the context of a case-control study, the validity of plasma and platelet fatty acid measurements will always be open to some question as it is possible that they may be affected by acute events, such as those due to infarction, or be the result of dietary or other behaviors that change after the diagnosis of the disease. This problem is avoided by using blood specimens that were collected before the occurrence of CHD. Thus, Miettinen and co-workers (1982) used serum that had been collected and stored for a cohort of 1222 men to measure fatty acid levels of 33 men who died of CHD and compared these levels with those from 64 men matched on the basis of hyperlipidemia and other risk factors and who remained free of CHD. A low content of linoleic and linolenic acids and high levels of palmitic and stearic acids (both saturated fats) in the serum phospholipids were found to be predictive of fatal CHD; however, these associations were not present for fatty acids measured in the cholesteryl ester or triglyceride fraction.

Biochemical analyses of blood or tissue fatty acids have the potential for providing information about diet that may be difficult to obtain from questionnaires as the type of fat used in prepared foods and the degree to which it has been modified by processing may be difficult for an individual to know. The studies noted previously, however, do not provide a clear or consistent picture of the the relation between fatty acid intake and risk of CHD. Although lower

levels of linoleic acid have been seen in several tissues of patients with coronary heart disease in some studies, no independent association was found in others. As noted by Wood and colleagues (1987), many of these studies were small, the cases and sources of controls were often inadequately described, and the effects of potential confounders, such as cigarette smoking, were frequently not reported. Disconcertingly, no independent association of adipose linoleic acid level with myocardial infarction was noted in the the study of Wood and coworkers (1987), which was one of the larger and better conducted investigations. Although improvements in the size and design of future studies should provide valuable and more consistent results, the interpretation of tissue fatty acid analyses is still subject to several limitations (see Chapter 9). First, these data provide an indication of the *proportion* of fatty acids in the specific tissues; thus the percentage of one fatty acid will be affected by changes in other fatty acids. For this reason, controlling for a measure of total fat intake, necessarily assessed by questionnaire, could enhance the interpretation of findings; however, this has never been done to date. In addition, tissue levels of fatty acids are affected by factors other than dietary intake, such as individual differences in absorption and metabolism as well as selective incorporation into tissues or blood lipid fractions. For example, intake of the major saturated fatty acids is of great interest, but this is apparently poorly reflected in tissue levels as these fatty acids are also synthesized endogenously from carbohydrates. For this reason, minor fatty acids that are not synthesized by humans may provide better markers of diet; however, this remains to be established. Again, simultaneous data on both dietary intake and tissue fatty acids would greatly enhance the interpretation of the biochemical analyses.

TRIALS OF DIET IN THE PREVENTION OF CORONARY HEART DISEASE

The most direct test of the diet–heart hypothesis is to conduct a randomized trial to determine whether reducing saturated fat and cholesterol intake or increasing polyunsaturated fat intake, or both, decreases the risk of CHD. Unfortunately, clear evidence from such studies has been elusive, largely due to practical difficulties in conducting such large scale trials.

Early trials of dietary prevention of CHD were conducted among institutionalized men, specifically residents of the Los Angeles Veterans Administration Hospital (Dayton et al., 1969) and two Finnish mental hospitals (Turpeinen et al., 1979). In both studies, patients passively received modified diets; cholesterol intake was reduced largely by a reduction in egg consumption and polyunsaturated fats were increased to approximately 20 percent of calories by substitution for saturated fats in many foods. In the U.S. Veterans study, 846 men were randomized to either of the above diets and followed for up to 8 years. Seventy-one control men and 54 men on the special diets developed definite myocardial infarction or sudden death; this difference was not statistically significant. When cerebral infarction and other secondary endpoints were included, however, 47.7 percent of the control group and 31.3 percent of the experimental group devel-

oped an event ($p = 0.02$). In the Finnish study, approximately 250 men in one hospital received the modified diet and a similar number in the other served as a control; after 6 years the diets were reversed for the two institutions. Coronary heart disease rates were reduced on the modified diet: 51 percent lower for CHD deaths alone ($p = 0.10$) and 67 percent lower for CHD deaths and major electrocardiogram change ($p = 0.001$). Although the hospitals were crossed over after 6 years, the interpretation of this study is limited by the fact that the unit of randomization was the institution rather than the individual, thus providing an effective randomized sample size of two. In neither study was a benefit for total mortality observed.

Although promising results were provided by the two studies of institutionalized patients, greater interest existed in demonstrating that the diets of free-living individuals could be modified to lower serum cholesterol sufficiently to reduce CHD rates. Therefore, the Diet–Heart Study was initiated to serve as a pilot study for a large-scale national trial (National Diet–Heart Study Research Group, 1968). Foods were modified and provided to participants in a double-blind manner. Over a 1-year period, serum cholesterol dropped about 11 percent in the modified diet group compared with the control group. Despite this successful reduction of serum cholesterol, a full-scale trial was never mounted because of the large number of subjects and accompanying costs required to detect a reduction in CHD.

To reduce the number of study subjects required, two large trials were mounted to evaluate the effect on CHD rates of dietary modification with simultaneous reduction of other risk factors. Such a study design can provide information that is applicable in practice, but may produce a result that is difficult to interpret if a benefit is found. This is because the interventions are completely confounded and reduced incidence could be due to only one, or any combination, of the interventions. In the Multiple Risk Factor Intervention Trial (MRFIT), over 12,000 U.S. men at high risk of CHD were randomly assigned to either an intensive program of dietary modification, smoking cessation, and blood pressure control (Special Intervention (SI) group) or annual check-up (Usual Care group) (Multiple Risk Factor Intervention Trial Research Group, 1982). Interventions for both smoking (45 percent quit rate) and blood pressure control (approximately 11 mmHg reduction in diastolic blood pressure) were both highly successful. The dietary intervention was not only less successful (serum cholesterol was reduced by 7.2 percent in the SI group), but a similar reduction also occurred in the control group. Thus, the difference between groups in serum cholesterol during the intervention period was only about 2 percent, clearly an insufficient contrast to test the diet–heart hypothesis. No significant reduction in CHD mortality was found between the groups during the 10-year follow-up period and the difficulty of distinguishing between the effects of simultaneous interventions was not encountered.

More convincing support for the diet–heart hypothesis was obtained from the Oslo Heart Study (Hjermann et al., 1981, 1986). In this trial, 1232 normotensive men with high serum cholesterol levels, 80 percent of whom also smoked, were randomly assigned to either a program of dietary intervention and smoking cessation, or to a control group. Men who were already following a

lipid-lowering diet, based on responses to a simple eight-item questionnaire, were excluded before randomization. Dietary intervention involved primarily a reduction in saturated fat and cholesterol; polyunsaturated fat intake was unchanged or increased only slightly. These changes were accompanied by an increase in total carbohydrates and fiber intake. During the intervention period, serum cholesterol fell by 17 percent compared with the control group after 1 year, and after 5 years, was 13 percent lower in the intervention group. The smoking intervention was less succesful; only 25 percent of the smokers in the intervention group stopped compared with 17 percent in the control group. After 5 years, the incidence of nonfatal myocardial infarction and fatal CHD was 47 percent lower in the intervention group compared with controls ($p = 0.03$). The authors conducted a series of multivariate analyses to assess the relative effects of dietary intervention (assessed by change in serum cholesterol) and smoking cessation, thus treating the data as an observational study rather than a randomized trial. As is apparent by the relatively ineffective impact on smoking cessation, they concluded that most of the reduction in CHD incidence in this trial could be attributable to reduction in serum cholesterol. Although active intervention stopped in this population after 5 years, a similar reduction (approximately 45 percent) was observed after 102 months and the difference in total mortality had become marginally significant (19 deaths in the intervention group compared with 31 in the control group).

The contrast between the unimpressive findings of the MRFIT and the clear results of the Oslo study, both in serum cholesterol reduction and effect on incidence of CHD, may be in part related to the selection of subjects. It has been pointed out by Caggiula and colleagues (1983) that the initial serum cholesterol levels of Oslo participants (325 mg%) were substantially higher than those of the MRFIT participants (254 mg%) and that for MRFIT participants with levels over 300 mg%, the decrease in the intervention group was 14 percent. Among the MRFIT men with high baseline serum cholesterol, however, the control group also decreased 11 percent. The overall differences in serum cholesterol between the two studies is notable, and the overall decline in serum cholesterol of the MRFIT control group clearly contributed to difficulty in creating a contrast between the groups. The reductions in serum cholesterol for both the MRFIT intervention and control participants with high levels on the first screening test, however, are suggestive of regression toward the mean, and are not directly comparable with the Oslo data, which used the mean of three baseline cholesterol determinations for their initial values. Perhaps more importantly, men who were already following a cholesterol lowering diet were excluded from the Oslo trial, whereas this was not true for the MRFIT study. As shown in Table 15–3, control participants in the Oslo study at 4 years after randomization (who, if anything, should have altered their diets in a favorable direction) had appreciably higher intakes of total and saturated fat and cholesterol than did MRFIT participants at baseline.

The intervention groups for the two studies, however, achieved very similar intakes of saturated and polyunsaturated fat and cholesterol during the trial. Given the recruitment methods for the MRFIT, participants were likely to be volunteers with above average health consciousness and "better" diets. Indeed,

Table 15–3. Dietary intake among a sample of control and intervention subjects in the Oslo Diet and Antismoking Trial and among the special intervention participants in the MRFIT study at baseline and during the intervention period

	Oslo study, 4-year data		MRFIT, special intervention group	
	Control (n = 23)	Intervention (n = 23)	Baseline (n = 12,847)	Intervention[a] (n = 5,308)
Total calories/day	2331	2248	2488	—
Total fat (% cal)	44.1	27.9	38.3	—
Saturated fat (% cal)	18.3	8.2	14.0	10
Polyunsaturated fat (% cal)	7.1	8.3	6.4	8.7
Cholesterol (mg/day)	527	289	451	265

[a]Mean values for years 1–3.

From Hjermann et al., 1981 and Caggiula et al., 1981.

the nutrient data obtained at baseline in the MRFIT study indicated that participants consumed less saturated fat and dietary cholesterol and more polyunsaturated fat than had been anticipated, so the potential for lowering serum cholesterol was reduced (Caggiula et al., 1983). Thus, the MRFIT study tended to enroll men with high serum cholesterol levels who were less likely to be responsive to diet and the Oslo study tended to exclude men whose high serum cholesterol levels would be unresponsive to diet. The finding that dietary intervention appeared to be more effective in the Oslo study should thus not be surprising.

This point illustrates an often over-looked issue in designing dietary intervention trials; subjects are frequently selected on the basis of an intermediary factor (e.g., serum cholesterol) or on the basis of high overall risk of a clinical outcome (such as CHD or cancer). It is generally better, however, to select a population susceptible to the intervention on the basis of the dietary factors that will be directly altered by the intervention, such as high saturated fat and cholesterol intake. Although the number of endpoints can be increased by selecting a high-risk population on the basis of other factors, such as genetic predisposition, the statistical power of a study can actually be reduced by including persons whose increased risk is not influenced by the intervention (Stampfer et al., 1988). In the case of CHD, for which serum cholesterol is an intermediary factor that is partly determined by diet, a maximally susceptible population can be identified on the basis of both diet and serum cholesterol. For other diseases, such as many cancers for which an intermediary factor has not been identified, baseline intake would appropriately be a greater part of the selection criteria in a dietary intervention study.

In addition to the trials noted earlier that have directly examined the effect of change in diet on the incidence of CHD, other relevant evidence has been provided by trials of serum cholesterol-lowering drugs. In the Lipid Research Clinics Coronary Primary Prevention Trial, 3806 men with elevated serum cholesterol were assigned at random to placebo or cholestyramine, a bile acid

sequestrant that effectively lowers serum cholesterol. The 7-year cumulative incidence of nonfatal myocardial infarction or fatal CHD was 7 percent in the cholestryamine group and 8.6 percent in the placebo group; this difference was statistically significant using a one-tailed test (Lipid Research Clinics Program, 1984). A more pronounced difference was found in a recent trial comparing gemfibrozil, another cholesterol-lowering drug, with placebo; after 5 years of treatment the incidence rate of CHD was 6.8 per 1000 person-years in the treatment group and 9.4 per 1000 in the control group. Although a causal relationship between LD–cholesterol and CHD was already established beyond reasonable doubt before the publication of these trials, they have clearly documented the effectiveness of intervention; reductions in serum cholesterol by adult men can lead to a lower incidence of CHD within a few years.

Due to the practical difficulties in conducting primary prevention trials of diet and heart disease among free-living humans, studies in animals play an important role. Although the relevance of many animal models to human CHD can be questioned, studies among non-human primates are likely to be particularly germane and do contribute support to the diet–heart hypothesis. For example, Wissler and Vesselinovitch (1975) compared the development of atherosclerotic lesions in rhesus monkeys fed an average American diet with lesions among monkeys fed a "prudent diet" that was lower in calories, saturated fat, and cholesterol. Although monkeys fed the usual American diet developed four times as many coronary atherosclerotic lesions as the prudently fed animals, it is difficult to determine whether this was related to the composition of the diet as the prudent diets were nearly 40 percent lower in total energy availability and the animals fed this diet must have been substantially leaner. More recently, Eggen and co-workers (1987) have demonstrated that diet-induced atherosclerotic lesions in rhesus monkeys can regress if cholesterol is removed from the diet.

In summary, intervention trials have contributed important support to the diet–heart hypotheses, even though the ideal randomized trial among a free-living population has not been and may never be, conducted. Nevertheless, the information provided by these dietary trials has major limitations. Even in the Oslo study, which provided the clearest findings, the effect of intervention was only marginally statistically significant, meaning that the confidence intervals necessarily include values of almost no effect to implausibly large effects. With the added problem that the findings are to some extent inextricably confounded with a modest reduction in cigarette smoking, these data cannot provide firm quantitative information on the effectiveness of the overall dietary intervention or the components of the dietary change. In addition, the findings of the Oslo study must be generalized with caution as the subjects had extremely high serum cholesterol levels by contemporary U.S. standards.

CRITIQUE OF THE CLASSIC DIET–HEART HYPOTHESIS

Despite decades of intensive scientific interest and scrutiny, the diet–heart hypothesis remains unproven by the most rigorous scientific standards; specifi-

cally, replicated randomized experiments (Ahrens, 1985). Nevertheless, a massive body of indirect evidence provides general support for the hypothesis. Serum cholesterol is clearly causally related to risk of CHD, and a reduction of serum cholesterol, which can be achieved by alterations in dietary lipids, generally reduces the incidence of CHD. Like most maturing hypotheses, however, the original formulation of the diet–heart hypothesis requires refinements. The next section discusses the inadequacies of the classic diet–heart hypothesis, indicates further questions that merit investigation, and suggests additional dietary hypotheses that are not mutually exclusive.

Inadequacies of the Classic Diet–Heart Hypothesis

One obvious oversimplification of the diet–heart hypothesis is that total serum cholesterol does not represent the complete effect of blood lipids on risk of CHD. Partly as a result of improved analytic methods, blood lipids have been fractionated into successively finer components. Total serum cholesterol has been subdivided by several methods, including ultracentrifugation, electrophoresis, and precipitation techniques. One of the most useful and well established distinctions has been between high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and very low-density lipoprotein cholesterol (VLDL-C); in contrast to the other fractions, HDL-C is strongly protective for CHD (Gordon, 1988). HDL-C has been further fractionated into HDL-2 and HDL-3, and LDL subtypes have been further defined on the basis of particle size; the epidemiologic importance of these additional subdivisions is less well established. The importance of more transient lipid fractions, such as VLDL-C and chylomicrons as measured in the postprandial state, remains to be defined.

A second oversimplification of the diet–heart hypothesis is that serum cholesterol represents the full effect of diet on blood lipids. Although dietary lipids clearly influence total serum cholesterol and LDL-C (the major component of total cholesterol), some aspects of diet, such as alcohol intake, have major effects on HDL-C but relatively small effects on total cholesterol. In addition, total energy balance, reflected in degree of adiposity, is more strongly related to HDL-C (inversely) and VLDL-C (positively) than to total serum cholesterol (Rhoads et al., 1976).

Accumulating evidence also indicates that the blood cholesterol responses of individuals differ substantially in response to changes in dietary lipids (Jacobs et al., 1983; Katan et al., 1986). For the same increase in dietary cholesterol or saturated fat, the cholesterol levels of most persons will increase, but some will remain essentially unchanged and a few will increase dramatically. These differences in response are reproducible within individuals, indicating that this variation in response is not simply random statistical variability. Similarly, it has long been noted that the mean cholesterol levels are frequently distinctly different for populations with markedly different rates of CHD, but that within each of these populations there exists a substantial degree of variability among persons. Factors that predict response to dietary changes are poorly defined at present, but are likely to be largely genetic. It is thus likely that a diet–heart hypoth-

esis must include a strong interaction with genetic predisposition to hypercholesterolemia and CHD.

A third limitation of the classic diet–heart hypothesis is that atherosclerosis and progressive coronary narrowing do not fully represent the pathophysiology of acute myocardial infarction. Indeed, the common clinical presentation of sudden, unheralded, catastrophic chest pain is not compatible with a slowly progressive process that develops over decades. With the advent of routine angiography performed at the time of infarction, it has become clear that acute thrombosis (to which atherosclerotic lesions may predispose) plays a central role in the development of myocardial infarction (Dewood et al., 1980; Rentrop et al., 1986). It is thus important to consider the possible effects of diet on factors that influence thrombosis. Higher fibrinogen levels increase the risk of CHD (Meade et al., 1986) and, in one study, dietary fiber intake has been associated with lower fibrinogen levels (Fehily et al., 1982). Platelet aggregability, which also alters the likelihood of thrombosis, clearly has a major influence on risk of CHD (Steering Committee of the Physicians' Health Study Research Group, 1988) and dietary factors may well act by this mechanism. For example, high intake of omega-6 polyunsaturated fatty acids may increase platelet aggregation (Renaud et al., 1986), whereas omega-3 fatty acids may be inhibitory (discussed later). The possibility that the effects of diet on risk of CHD are not entirely mediated by total serum cholesterol was supported by the findings of Shekelle and co-workers (1981), who found an association between lipid intake and CHD mortality after controlling for serum cholesterol in a multivariate analysis. Thus, a second general diet–heart hypothesis must be considered whereby aspects of the diet, which may be multiple, act through one or more pathways to influence the probability of thrombosis formation in the coronary arteries.

Although the classic diet–heart hypothesis has proved useful and is supported by numerous lines of evidence, it is not likely to completely describe the relation of diet with CHD. Due to the complexities noted earlier, the effect of diet on risk of CHD cannot be completely or reliably predicted solely on the basis of the response of serum total cholesterol. Therefore, studies that examine directly the relation of dietary factors with incidence of CHD are important.

Further Questions Regarding the Diet–Heart Hypothesis

Independent Effects of Dietary Saturated Fat, Cholesterol, and Polyunsaturated Fat

Even if it is accepted that diets high in saturated fat and cholesterol and low in polyunsaturated fat increase the risk of CHD, knowledge of the independent effects of these three components is of great practical importance as they can be manipulated separately in diets. For example, cholesterol intake is greatly influenced by changes in egg consumption without major changes in saturated or polyunsaturated fat intake. Substantial increases in polyunsaturated fat intake can be made by the substitution of vegetable oils for either other fats or carbohydrates, which may or may not change intake of saturated fat or cholesterol. Available data do not provide a clear indication of the independent quantitative effects of these three lipid components. The optimal level of polyunsaturated fat

intake is particularly unclear. Based simply on the relationship with total serum cholesterol, one would conclude that maximizing the intake would be desirable, and it has been suggested that our current intake be increased, while simultaneously decreasing saturated fat, to attain a ratio of polyunsaturated to saturated fats of 1.0 (Consensus Conference, 1985). Using platelet aggregability rather than serum cholesterol to evaluate the effect of diet, Renaud and colleagues (1986) have suggested that a dietary P/S ratio of 0.6 to 0.8 may be superior to a ratio of 1.0. Because the dose-response relationship between polyunsaturated fat intake and risk of CHD has not been determined directly, the optimal intake remains uncertain.

Fatty Acid Structure

Saturated and polyunsaturated fats have frequently been treated as specific entities, however, the importance of their heterogeneity is becoming increasingly recognized. For example, the physiologic effects of omega-6 (primarily represented in the diet by linoleic acid) and omega-3 fatty acids (represented largely by linolenic acid and longer chain fatty acids derived from marine sources) are sometimes opposing (Leaf and Weber, 1988). The possibility that these types of polyunsaturated fats may compete with each other at enzymatic sites suggests that their interaction needs to be considered in epidemiologic studies. In addition, not all saturated fats have a similar effect on serum cholesterol; stearic acid, one of the dominant saturated fats in chocolate, does not appear to raise LDL-C in the same manner as does palmitic acid, a more common saturated fat in most animal products (Bonahome and Grundy, 1988). Short-chain saturated fats, such as those found in coconut fat, appear to be more atherogenic in animals than predicted solely on the basis of their relationship with serum cholesterol (Wissler and Vesselinovitch, 1975). A substantial portion of the dietary fat in the American diet consists of synthetic *trans*-isomers of natural fatty acids produced in the hydrogenation of liquid vegetable oils, whereby these are converted to margarine and vegetable shortening. The long-term effects of these isomers have been little studied in humans; in one case-control study, British men with CHD had higher levels of partially hydrogenated fatty acids in their subcutaneous fat compared with controls (Thomas and Scott, 1981). More detailed studies of specific fatty acid intake in relation to risk of CHD are clearly warranted.

Nutrients and Foods

Investigations and discussions on diet and CHD have largely focused on nutrients rather than on foods. Although this has scientific justification, practicality dictates that intake of foods should also be examined as nutrients are not usually consumed in pure chemical form. For example, good reasons exist to hypothesize that high consumption of eggs increases the risk of CHD, due to their high cholesterol content. This cannot, however, be concluded with certainty as eggs contain a wide variety of other substances that could conceivably counterbalance an adverse effect of the cholesterol. Few studies have examined the relation of egg intake with CHD incidence and none have provided clear evidence of an association. Furthermore, findings of an association between

intake of foods or groups of foods and risk of CHD may provide guidance for selecting diets, even when the responsible nutrient or substance is not yet identified. As will be discussed later, fiber intake has been reported to be inversely related to risk of CHD in several studies. Although interesting, further information on the fiber-containing foods that contributed to this protective effect would be extremely helpful, because not all forms or sources of fiber have the same physiologic effects. This type of information was provided by Morris and colleagues (1977), who found that only fiber from grain sources contributed to a reduced risk of CHD. In addition, the possibility remains that factors other than fiber in these foods, such as magnesium, are responsible for the reduced incidence of disease; the identification of these foods may provide the basis of alternate hypotheses that can be tested in other data. Thus, the investigation of foods in relation to occurrence of CHD is likely to further our basic understanding of this disease as well as provide practical guidance for eating.

Additional Diet–Heart Hypotheses

Although substantial indirect evidence supports the classic diet–heart hypothesis, the magnitude of any association is likely to be modest for ranges of diet attainable within Western culture, leaving open the possibility that other dietary factors may also influence the occurrence of CHD. Indeed, it is quite possible that intakes of saturated and polyunsaturated fats and cholesterol are less important than other dietary factors in determining risk of CHD within the United States or similar populations. In addition to acting through total serum cholesterol, effects of other dietary factors could be mediated by other lipid fractions, platelet aggregability, fibrinogen levels, blood pressure, glucose tolerance, endothelial damage, insulin resistance, obesity, and other mechanisms that have yet to be discovered. In this section some additional hypotheses relating dietary factors to CHD are briefly discussed. The objective is not to list exhaustively all possible hypotheses or explore them in depth, but rather to provide a sense of the scope of other factors that deserve consideration.

Obesity

Energy balance, as reflected by degree of adiposity, represents an important influence of diet on risk of coronary heart disease that is beyond the scope of this chapter, but that may well be more important than the composition of the diet (Manson et al., 1987). The effect of obesity is mediated through numerous mechanisms including hypertension, hyperglycemia, reduced HDL-C, and only to a slight extent through total serum cholesterol (Rhoads et al., 1976).

Alcoholic Beverages

Among the various dietary factors that have been examined in epidemiologic studies of CHD, the most consistent association has been an inverse relationship with moderate alcohol intake. This relationship has been reviewed elsewhere (Hennekens, 1983; Moore and Pearson, 1986; Stampfer et al., 1988); in the summary by Moore and Pearson, a protective relationship was observed in seven out of eight case-control and 12 of 14 cohort studies. In general, consumption

of one or two drinks of beer, wine, or liquor per day has corresponded to a reduction in risk of approximately 30 to 70 percent. Alcohol increases blood levels of HDL-C, which is likely to be the most important mechanism underlying its protective effect. In one study, however, an inverse relationship with CHD existed even after controlling for level of HDL-C, suggesting that additional causal pathways might be involved (Criqui et al., 1987). At higher consumption levels, the risk of death attributed to CHD has been found to be similar to or even higher than the risk among nondrinkers (Moore and Pearson, 1986), which may be due to direct myocardial toxicity or the tendency for alcohol to induce arrhythmias, resulting in the assignment of myocardial infarction as a cause of death.

Despite the large body of evidence indicating that moderate alcohol intake is associated with a reduced risk of CHD, some have been reluctant to accept this as a causal relationship. One argument has been that nondrinkers include a substantial number of former alcoholics who continue to be at excess risk of CHD, thus creating the perception of elevated risk among this group relative to moderate drinkers. This hypothesis, however, was not supported by the findings of Rosenberg and colleagues (1981), who found a reduction in risk among moderate drinkers relative to lifetime nondrinkers.

A second argument against the hypothesis of reduced risk has been that alcohol intake primarily causes an elevation in levels of HDL-3 fraction, whereas it is HDL-2 rather than HDL-3 that is related to reduced risk of CHD. In general, the use of a hypothetical mechanism as evidence against a large body of empirical data provides a very weak argument. In this instance, the data indicating that alcohol exclusively affects HDL-3 are based on a very small study (Haskell et al., 1984) contradicted by other findings (Carmago et al., 1985; Valimaki et al., 1986). Furthermore, a protective relationship has been observed between HDL-3 and risk of CHD in several studies (Hennekens et al., 1987; Miller, 1987). Even if HDL-3 was not a primary intervening variable, it would not mean that alcohol could not be reducing CHD by other mechanisms.

Fiber

A growing literature suggests that certain types of fiber in the diet can reduce hyperglycemia (Rivellese et al., 1980; Potter et al., 1981) and exert a beneficial effect on blood lipids (Behall et al., 1984; Anderson et al., 1984). The relationship between fiber intake and risk of CHD has been reported in four epidemiologic studies and an inverse relationship has been noted in each (Table 15–2). Morris and colleagues (1977) found that men in the highest third of dietary fiber consumption experienced about one third the risk of CHD during the next 20 years. These authors recognized that this relationship was partly explained by the tendency for men with low total energy intake to have an elevated risk of CHD (as a result of reduced physical activity); however, even when expressed as a nutrient density (grams/1000 kcal), the inverse relationship with fiber intake remained statistically significant. The protective effect was entirely attributable to fiber from grains; fiber from fruits, vegetables, and other sources was not associated with CHD incidence. Kushi and co-workers (1985), in a follow-up study of Irish siblings, found that those in the highest third of fiber consumption had about only about half the risk of CHD death compared with those in the lowest

third. This finding persisted after adjustment for other risk factors including total serum cholesterol and blood pressure. When adjusted for total energy intake, however, this relationship was no longer statistically significant (Kushi, 1987). A similar inverse relationship was seen in the follow-up study of Dutch men conducted by Kromhout and colleagues (1982). Khaw and Barrett-Connor (1987), in a study based on a 12-year follow-up after obtaining a single 24-hour recall, found an inverse relationship between dietary fiber intake and CHD mortality among both men and women. Overall, the relative risk for a 6-g increase in dietary fiber, corresponding to about one standard deviation, was 0.74 (95% confidence interval 0.58 to 0.94).

Although each of these studies of dietary fiber has been quite small the consistency of the findings is remarkable. Nevertheless, a cautious interpretation of the data would suggest that some aspect of plant products, not necessarily a specific type of fiber, may reduce the risk of CHD. This relationship certainly deserves further examination, with focus on types and sources of fiber as well as a full consideration of other factors in fruits and vegetables.

Fish and Omega-3 Fatty Acids

Low rates of CHD in Japan and Greenland have led to speculation that the high consumption of fish in these areas might be protective (Bang et al., 1980; Kagawa et al., 1982). This hypothesis was supported by the finding of Kromhout and co-workers (1985), that Dutch men consuming more than 30 g of fish per day had only about half the risk of fatal CHD compared with men who consumed none. Lower rates of CHD mortality among men and women who consumed higher amounts of fish were also observed in prospective studies reported by Shekelle and colleagues (1985) and Norell and colleagues (1986), but not in two large studies conducted by Vollset and co-workers (1985) among Norwegian men or by Curb and Reed (1985) among Japanese men living in Hawaii.

The hypothesis that fish intake may reduce the occurrence of CHD receives considerable support from rapidly accumulating evidence that omega-3 fatty acids, which are predominantly provided by fish and other marine animals, have a wide variety of presumably favorable physiologic effects. These effects, recently reviewed by Leaf and Weber (1988), include a potent reduction in VLDL, inhibition of thromboxane production and increase in prostacyclin synthesis with a resulting reduction in thrombotic tendency, reduction in blood viscosity, increase in fibrinolytic activity, and, perhaps, reduction in blood pressure.

Measurements of long-chain (20 and 22-carbon) omega-3 fatty acids in blood and tissue can provide markers of past fish intake, providing another method to examine the relation between fish intake and risk of CHD. In two small case-control studies, each including fewer than 30 patients with CHD, no significant associations were reported between risk of CHD and long-chain omega-3 fatty acid levels in adipose tissue or platelets (Wood et al., 1984) or in specific serum lipid fractions (Schrade, 1961). Using frozen sera from a prospective study, Miettinen and colleagues (1982) found that phospholipid omega-3 fatty acid levels tended to be lower among the 33 men who subsequently developed myocardial infarctions than among a series of 64 men who remained free of CHD. The reverse was true, however, for omega-3 fatty acid levels in the triglyceride frac-

tion. Wood and colleagues (1987) measured adipose tissue and platelet fatty acid levels in 80 men with acute myocardial infarction, 108 men with angina pectoris, and 391 men without CHD. In both adipose and platelets, the long-chain omega-3 fatty acid levels were consistently lower for patients with CHD and for several comparisons these differences were statistically significant. The interpretation of findings based on tissue measurements of long-chain omega-3 fatty acids must be tempered by the findings that some endogenous synthesis may occur from linolenic acid and that omega-6 fatty acids, such as linoleic acid, may competitively inhibit this synthesis. Because linoleic acid levels have, if anything, tended to be lower among cases compared with controls, such competition is unlikely to explain the lower levels of omega-3 fatty acids among cases found in the largest of these studies (Wood et al., 1987).

Although the possibility that increased intake of fish or supplements of omega-3 fatty acids may reduce the incidence of CHD appears promising, existing data are incomplete and not entirely consistent. Studies that include both measures of fish consumption and measurements of tissue fatty acid levels would be particularly useful. To avoid the possibility that events surrounding the diagnosis of CHD or recent changes in diet might affect blood levels of fatty acids, either prospective studies based on prediagnostic blood specimens or the use of adipose for case-control studies would be most desirable. Because details of any dose-response relationship are of particular interest, such studies must be larger than the existing studies using biochemical measurements.

Animal Protein

In numerous laboratory experiments, animal protein (usually beef protein or casein) has induced hypercholesterolemia and accelerated atherogenesis (Munro et al., 1965; Vahouny et al., 1984). The effects of protein and type of protein in humans are less clear; whereas Sirtori and colleagues (1979) found that substitution of vegetable for animal proteins reduced serum lipids, such an effect was not observed by Sacks and co-workers (1983). Potter and colleagues (1980, 1981) have suggested that saponins contained in some soybean protein preparations, and which sequester bile acids in the colon, may be responsible for the hypocholesterolemic effects seen in some experiments that attributed this effect to the amino acid composition of the vegetable product. The relationship of type of protein with risk of CHD has received little attention in epidemiologic studies.

Complex and Simple Carbohydrates

In several studies total carbohydrate intake has been inversely related to risk of CHD (Gordon et al., 1981; McGee et al., 1984), although in the Puerto Rican data it is not clear that this effect was independent of total caloric intake. Because total carbohydrate tends to vary inversely with fat intake if calories are held constant, the relation of total carbohydrate intake with CHD is practically inseparable from the classic diet–heart hypothesis. For this reason, the type of carbohydrate may be of more interest than the total intake. Yudkin (1963) has proposed that sucrose is a major determinant of CHD, largely based on international comparisons. Although this hypothesis has been questioned (McGandy et al., 1967), it has not been securely confirmed or refuted. Sucrose may decrease

blood HDL-C levels (Morrison et al., 1980), have an adverse effect on glucose tolerance (Reiser et al., 1981), and, along with other carbohydrates, elevate blood pressure and catecholamine excretion in animals (Young and Lansberg, 1981) and possibly humans (Rowe et al., 1981). In monkeys, a fructose-rich diet appeared to be atherogenic (Kritchevsky et al., 1974).

Selenium

Selenium is an essential trace element and its deficiency has been associated with Keshan disease, a cardiomyopathy described in China. Puska and colleagues (1982) reported an association between low selenium levels in prospectively collected blood and subsequent risk of CHD in Finland, a finding that has been confirmed and extended to stroke in a larger study from the same country (Virtamo et al., 1985). In a small study from southern Finland, it was suggested that this inverse association may result from confounding by intake of fish, which, in Finland, is major source of both long-chain omega-3 fatty acids and selenium (Miettinen et al., 1983). Although fish is a relatively less important source of selenium in the United States, this study underlines the need for simultaneous measurements of multiple nutritional factors in epidemiologic studies. In one U.S. study, an inverse association between blood selenium levels and degree of atherosclerosis has been found among patients undergoing coronary artery angiography (More et al., 1984).

Vitamins A, C, and E

Vitamin E has long attracted popular attention as a potential agent to reduce CHD symptoms; however, in short-term studies it has not clearly reduced symptoms due to angina pectoris (Rinzler et al., 1985; Anderson and Reid, 1974). In a poorly controlled trial, patients with intermittent claudication treated with vitamin E were said to improve (Haeger, 1973). Roberts (1981) has suggested that vitamin E may increase the risk of thrombotic disease; although his study was poorly controlled, a convincing refutation has not been produced.

A time trend analysis of data from Israel has suggested an inverse association between vitamin A intake and CHD rates (Palgi, 1981). An antioxidant hypothesis, analogous to that suggested for cancer, has been proposed for CHD in which multiple nutrients that function as antioxidants, such as vitamin E, beta-carotene, and vitamin C, or components of antioxidative enzymes, such as selenium or copper, act jointly to protect the arterial intima from degradation or "aging" (Harmon, 1982; Gey, 1986). Gey and colleagues have reported ecologic data based on blood samples and national CHD rates in Europe suggesting that the combination of low plasma alpha-tocopherol and ascorbate are associated with high CHD rates (Gey, 1986). In a study among patients being evaluated by angiography, the degree of coronary occlusion was inversely related to blood leukocyte ascorbate levels among both smokers and nonsmokers (Ramirazan and Flowers, 1980). Such data are provocative, but need confirmation.

Vitamin B_6

More than 40 years ago Rinehart and Greenberg (1951) demonstrated that low vitamin B_6 intake produced arterial intimal damage in monkeys. Although an

inverse association between low plasma B_6 levels and CHD has been reported in two small case-control studies (Gvozdova et al., 1966; Vermaak et al., 1986), relevant epidemiologic data are limited (Willett, 1985).

Calcium and Magnesium

A large and inconclusive literature has examined the relationship between water hardness, primarily although not uniquely characterized by high calcium and magnesium levels, and CHD (Masironi et al., 1972). Although this is a weak way of evaluating the effects of these minerals as most individuals obtain them primarily through food rather than water, these relationships have apparently not been examined directly in case-control or cohort studies. A possible association between calcium intake and blood pressure (McCarron et al., 1984) has suggested one potential mechanism. Adequate magnesium intake is necessary to prevent CHD in animals (Seelig and Heggtveit, 1974; Seelig, 1980) and to maintain electrical stability in animals and humans (Karppanen, 1981; Luoma et al., 1983). In two autopsy series myocardial magnesium levels were lower among those dying of CHD compared with other causes (Chippenfield, 1978; Elwood and Beasley, 1981).

Other Minerals

Internationally, countries with higher chromium intake have lower rates of CHD (Anderson, 1981) and in a small case-control study, hair levels of chromium were inversely associated with risk of CHD; the relative risk for extreme quartiles was 6.4 (Cote et al., 1979). Chromium may reduce CHD risk by improving glucose tolerance, reducing total serum cholesterol, and increasing HDL cholesterol (Levine et al., 1968; Schroeder et al., 1970; Riales and Albrink, 1981; Anderson et al., 1983), although the latter effect did not always attain statistical significance (Anderson et al., 1983). One problem in addressing relationships of chromium with disease is that blood levels are extremely low and easily contaminated. Keratinous tissues, however, concentrate this element and supplements have been shown to raise levels in hair (Anderson et al., 1983).

Klevay has hypothesized that low intake of copper, particularly when coupled with high zinc intake, increases the risk of CHD (Klevay, 1975, 1983). Although speculation has existed for some time that these minerals may influence CHD risk, there are few prospective data relating either their intake or tissue levels to CHD risk in humans. Based on prospectively collected sera, Kok and colleagues (1988) recently reported that higher levels of copper were associated with a 3.5-fold increased risk of fatal coronary heart disease, and that zinc levels had no clear relationship with disease. It remains unclear, however, whether these associations represent differences in intake or in metabolism or even chance, as the study included only 62 cases.

SUMMARY

Abundant indirect evidence supports the hypothesis that dietary lipids are a causal determinant of coronary heart disease. Nevertheless, the dose-response

relationships between specific fatty acids and cholesterol intake and rates of CHD are not clearly defined. Furthermore, aspects of diet other than lipids may also be related to CHD; substantial evidence indicates that moderate alcohol intake reduces risk and other data suggest that some aspects of plants, perhaps a specific type of fiber, are associated with lower rates. A number of other hypotheses based on animal studies and theoretical mechanisms remain to be seriously investigated in human populations. Finally, data relating intake of specific foods to risk of CHD are extremely limited. Further epidemiologic data are needed to provide sound dietary guidance for persons wishing to reduce their risk of CHD.

REFERENCES

Ahrens, E. H. (1985). The diet-heart question in 1985: Has it really been settled? *Lancet i,* 1085–1087.

American Heart Association (1984). Recommendations for treatment of hyperlipidemia in adults: A joint statement of the Nutrition Committee and the Council on Atherosclerosis. *Circulation 69,* 1067A–1090A.

Anderson, J. W., L. Story, B. Sieling, et al. (1984). Hypocholesterolemic effects of oat-bran or bean intake for hypercholesterolemic men. *Am. J. Clin. Nutr. 40,* 1146–1155.

Anderson, R. A. (1981). Nutritional role of chromium. *Sci. Total Environ. 17,* 13–29.

Anderson, R. A., M. M. Polansky, N. A. Bryden, K. Y. Patterson, C. Veillon, and W. H. Glinsmann (1983). Effects of chromium supplementation on urinary Cr excretion of human subjects and correlation of Cr excretion with selected clinical parameters. *J. Nutr. 113,* 276–281.

Anderson, T. W. (1979). The male epidemic. In *Proceedings of the conference on the decline in coronary heart disease mortality.* (R. J. Havlik, M. Feinleib, eds.) U.S. Dept. Health, Education, and Welfare, Public Health Service, NIH, Washington, D.C. (NIH Publ. No. 79-1610), pp. 42–47.

Anderson, T. W., and D. B. Reid (1974). A double-blind trial of vitamin E in angina pectoris. *Am. J. Clin. Nutr. 27,* 1174–1178.

Anitschkow, N. (1967). A history of experimentation on arteriosclerosis in animals. In *Cowdry's Arteriosclerosis* (H. T. Blumenthal, ed.) 2nd ed., Macmillan, Springfield, pp. 21–44.

Bang, H. O., J. Dyerberg, and H. M. Sinclair (1980). The composition of the Eskimo food in north western Greenland. *Am. J. Clin. Nutr. 33,* 2657–2661.

Behall, K. M., K. H. Lee, and P. B. Moser (1984). Blood lipids and lipoproteins in adult men fed four refined fibers. *Am. J. Clin. Nutr. 39,* 209–214.

Bonanome, A. and S. M. Grundy (1986). Effect of dietary stearic acid on plasma cholesterol and lipoprotein levels. *N. Engl. J. Med. 318,* 1244–1248.

Burr, M. L. and P. M. Sweetnam (1982). Vegetarianism, dietary fiber, and mortality. *Am. J. Clin. Nutr. 36,* 873–877.

Caggiula, A. W., G. Christakis, M. Farraud, S. B. Hulley, R. Johnson, N. L. Lasser, J. Stamler, and G. Widdowson (1981). The Multiple Risk Factor Intervention Trial (MRFIT)—IV. Intervention on blood lipids. *Prev. Med. 10,* 443–475.

Caggiula, A. W., T. J. Orchard, and L. H. Kuller (1983). Epidemiologic studies of nutrition and heart disease. In E. B. Feldman, ed.: *Nutrition and Heart Disease.* New York: Churchill, Livingstone, pp. 1–27.

Camargo, C. A. Jr., P. T. Williams, K. M. Vranizan, J. J. Albers, and P. D. Wood (1985).

The effect of moderate alcohol intake on serum apolipoproteins A-1 and A-11: A controlled study. *J.A.M.A. 253,* 2854–2857.

Chipperfield, B. and J. R. Chipperfield (1978). Differences in metal content of the heart muscle in death from ischemic heart disease. *Am. Heart J. 95,* 732–737.

Connor, W. E., M. T. Cerqueira, R. W. Connor, R. B. Wallace, M. R. Malinow, and H. R. Casdorph (1978). The plasma lipids, lipoproteins, and diet of the Tarahumara Indians of Mexico. *Am. J. Clin. Nutr. 31,* 1131–1142.

Consensus Conference (1985). Lowering blood cholesterol to prevent heart disease. *J.A.M.A. 253,* 2080–2086.

Cote, M., L. Munan, M. Gagne-Billon, et al. (1979). Hair chromium concentration and arteriosclerotic heart disease. In D. Shapcott and J. Hubert, eds: *Chromium in Nutrition and Metabolism.* New York: Elsevier/North Holland Biomedical Press, pp. 223–228.

Criqui, M. H., L. D. Cowan, H. A. Tyroler, S. Bangdiwala, G. Heiss, R. B. Wallace, and R. Cohn (1987). Lipoproteins as mediators for the effects of alcohol consumption and cigarette smoking on cardiovascular mortality: results from the Lipid Research Clinics Follow-up Study. *Am. J. Epidemiol. 126,* 629–637.

Curb, J. D. and D. M. Reed (1985). Fish consumption and mortality from coronary heart disease (letter). *N. Engl. J. Med. 313,* 821.

Dawber, T. R., R. J. Nickerson, F. N. Brand, and J. Pool (1982). Eggs, serum cholesterol, and coronary heart disease. *Am. J. Clin. Nutr. 36,* 617–625.

Dayton, S., M. L. Pearce, S. Hashimoto, W. J. Dixon, and U. Tomiyasu (1969). A controlled clinical trial of a diet high in unsaturated fat in preventing complications of atherosclerosis. *Circulation 40* (Suppl II–1).

Dewood, M. A., J. Spores, R. Notske, L. T. Mouser, R. Burroughs, M. S. Golden, and H. T. Lang (1980). Prevalence of total coronary occlusion during the early hours of transmural myocardial infarction. *N. Engl. J. Med. 303,* 897–902.

Eggen, D. A., J. P. Strong, W. P. Newman, G. T. Malcom, and C. Restrepo (1987). Regression of experimental atherosclerotic lesions in rhesus monkeys consuming a high saturated fat diet. *Arteriosclerosis 7,* 125–134.

Elwood, P. and W. Beasley (1981). Myocardial magnesium and ischemic heart disease. *Artery 9,* 200–204.

Eskin, F. (1971). The role of the egg as a factor in the aetiology of coronary heart disease. *Community Health 2,* 179–184.

Fehily, A. M., J. E. Milbank, J. W. Yarnell, T. M. Hayes, A. J. Kubiki, and R. D. Eastham (1982). Dietary determinants of lipoproteins, total cholesterol, viscosity, fibrinogen, and blood pressure. *Am. J. Clin. Nutr. 36,* 890–896.

Finegan, A., N. Hickey, B. Maurer, and R. Mulcahy (1968). Diet and coronary heart disease: dietary analysis on 100 male patients. *Am. J. Clin. Nutr. 24,* 143–148.

Frick, M. H., O. Elo, K. Haapa, et al. (1987). Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia: safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N. Engl. J. Med. 317,* 1237–1245.

Friend, B. (1967). Nutrients in the United States Food Supply: a review of trends, 1909–1913 to 1965. *Am. J. Clin. Nutr. 20,* 907–914.

Garcia-Palmieri, M. R., P. D. Sorlie, and J. Tillotsen, et al. (1980). Relation of dietary intake to subsequent coronary heart disease incidence: The Puerto Rican Heart Health Program. *Am. J. Clin. Nutr. 33,* 1818–1827.

Gey, K. F. (1986). On the antioxidant hypothesis with regard to arteriosclerosis. In J. C. Somogyi, ed.: *Scientific Evidence for Dietary Targets in Europe Bibliotheca Nutr. Dieta Vol. 37.* Basel: Karger.

Gordon, T. (1988). The diet-heart idea: outline of a history. *Am. J. Epidemiol. 127,* 220–225.

Gordon, T., A. Kagan, M. Garcia-Palmieri, W. B. Kannel, W. J. Zukel, J. Tillotson, P. Sorlie, and M. Hjortland (1981). Diet and its relation to coronary heart disease and death in three populations. *Circulation 63,* 500–515.

Grundy, S. M., D. Bilheimer, H. Blackburn, W. V. Brown, P. O. Kwitervich Jr., F. Mattson, G. Schonfeld, and W. H. Weidman (1982). Rationale of the Diet-Heart Statement of the American Heart Association. Report of Nutrition Committee. *Circulation 65,* 839A–854A.

Gvozdova, L. G., E. G. Paramonova, E. V. Goryachenkova, and L. A. Polyakova (1966). The level of pyridoxal co-enzymes in the plasma of patients with coronary atherosclerosis kept on a curative diet and after additional intake of vitamin B6. *Vop. Pitan. 25,* 40–44.

Haeger, K. (1973). Walking distance and arterial flow during long term treatment of intermittent claudication with d-alpha-tocopheral. *Vasa 3,* 280–287.

Harmon, D. (1982). In W. A. Pryor, ed.: *Free Radials in Biology* New York: Academic Press, pp. 255–275.

Haskell, W. L., C. Camargo Jr., P. T. Williams, K. M. Vranizan, R. M. Krauss, F. T. Lindgren, and P. D. Wood (1984). The effect of cessation and resumption of moderate alcohol intake on serum high-density-lipoprotein subfractions. A controlled study. *N. Engl. J. Med. 310,* 805–810.

Heady, J. A. (1961). Diets of bank clerks. Development of a method of classifying the diets of individuals for use in epidemiologic studies. *J. R. Stat. Soc. 124,* 336–361.

Hegsted, D. M., R. B. McGandy, M. L. Myers, and F. J. Stare (1965). Quantitative effects of dietary fat on serum cholesterol in man. *Am. J. Clin. Nutr. 17,* 281–295.

Hegsted, D. M. (1986). Serum-cholesterol response to dietary cholesterol: A re-evaluation. *Am. J. Clin. Nutr. 44,* 299–305.

Hennekens, C. H. (1983). Alcohol. In J. Stamler and N. Kaplan, eds.: *Prevention of Coronary Heart Disease: Practical Management of the Risk Factors.* Philadelphia: W. B. Saunders, pp. 130–139.

Hennekens, C. H., J. E. Buring, G. T. O'Connor, S. Z. Goldhaber, M. J. Stampfer, J. Breslow, and W. C. Willett (1987). Moderate alcohol consumption and risk of myocardial infarction. *Circulation (Suppl) 76,* IV–105.

Hjermann, I., K. Velve Byre, I. Holme, and P. Leren (1981). Effect of diet and smoking on the incidence of coronary heart disease. Report from the Oslo Study Group of a randomised trial in healthy men. *Lancet ii,* 1303–1310.

Hjermann, I., I. Holme, and P. Lerren (1986). Oslo Study Diet and Anti-smoking Trial. Results after 102 months. *Am. J. Med. 80,* 7–11.

Jacobs, D. R. Jr., J. T. Anderson, and H. Blackburn (1979). Diet and serum cholesterol: do zero correlations negate the relationship? *Am. J. Epidemiol. 110,* 77–87.

Jacobs, D. R., Anderson, J. T., P. Hannan, A. Keys, and H. Blackburn (1983). Variability in individual serum cholesterol response to change in diet. *Arteriosclerosis 3,* 349–356.

Kagawa, Y., M. Nishizawa, M. Suzuki, et al. (1982). Eicosapolyenoic acids of serum lipids of Japanese islanders with low incidence of cardiovascular diseases. *J. Nutr. Sci. Vitaminol. 28,* 441–453.

Kahn, H. A. (1970). Change in serum cholesterol associated with changes in the United States civilian diet 1909–1965. *Am. J. Clin. Nutr. 23,* 879–882.

Karppanen, H. (1981). Epidemiological studies on the relationship between magnesium intake and cardiovascular diseases. *Artery 9,* 190–199.

Katan, M. B., A. C. Beynen, J. H. deVries, and A. Nobels (1986). Existence of consistent

hypo- and hyper-responders to dietary cholesterol in man. *Am. J. Epidemiol. 123,* 221–234.

Kato, H., J. Tillotson, M. Z. Nichamen, G. G. Rhoads, and H. B. Hamilton (1973). Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: serum lipids and diet. *Am. J. Epidemiol. 97,* 372–385.

Katz, L. N. and J. S. Stamler (1953). *Experimental Atheroschlerosis.* Springfield: Charles C Thomas Publ., pp. 24–32.

Keys, A., J. T. Anderson, and F. Grande (1965). Serum cholesterol response to changes in the diet. *Metabolism 14,* 747–758.

Keys, A. and R. W. Parlin (1966). Serum-cholesterol response to changes in dietary lipids. *Am. J. Clin. Nutr. 19,* 175–181.

Keys, A. (1980). *Seven Countries: A multivariate analysis of death and coronary heart disease.* Cambridge: Harvard University Press, p. 252.

Keys, A. (1984). Serum cholesterol response to dietary cholesterol. *Am. J. Clin. Nutr. 40,* 351–359.

Khaw, K. T. and E. Barrett-Connor (1987). Dietary fiber and reduced ischemic heart disease mortality rates in men and women: A 12-year prospective study. *Am. J. Epidemiol. 126,* 1093–1102.

Kirkeby, K., S. Nitter-Hauge, and I. Bjerkedal (1972a). Fatty acid composition of adipose tissue in male Norwegians with myocardial infarction. *Acta Med. Scand. 191,* 321–24.

Kirkeby, K., P. Ingvalsen, and I. Bjerkedal (1972b). Fatty acid composition of serum lipids in men with myocardial infarction. *Acta Med. Scand. 192,* 513–519.

Klevay, L. (1975). Coronary heart disease: the zinc/copper hypothesis. *Am. J. Clin. Nutr. 28,* 764–774.

Klevay, L. (1983). Copper and ischemic heart disease. *Biol. Trace Elem. Res. 5,* 245–255.

Kok, F. J., C. M. Van Duijn, A. Hofman, G. B. Van der Voet, F. A. De Wolff, C. H. Paays, and H. A. Valkenburg (1988). Serum copper and zinc and the risk of death from cancer and cardiovascular disease. *Am. J. Epidemiol. 128,* 352–359.

Kritchevsky, D., L. M. Davidson, J. J. van der Watt, P. A. Winter, and I. Borsohn (1974). Hypercholesterolemia and atherosclerosis induced in vervet monkeys by cholesterol-free semisynthetic diet. *S. Afr. Med. J. 48,* 2413–2414.

Kromhout, D., E. B. Bosschieter, and C. de Lezenne Coulander (1965). The inverse relation between fish consumption and 20-year mortality from coronary heart disease. *N. Engl. J. Med. 312,* 1205–1209.

Kromhout, D., E. B. Bosschieter and C. de Lezenne Coulander (1982). Dietary fiber and 10-year mortality from coronary heart disease, cancer and all causes: The Zutphen Study. *Lancet ii,* 518–521.

Kromhout, D. and C. de Lezenne Coulander (1984). Diet, prevalence and 10-year mortality from coronary heart disease in 871 middle-aged men: The Zutphen Study. *Am. J. Epidemiol. 119,* 733–741.

Kushi, L. H. (1987). Total energy intake: Implication for epidemiologic analyses (letter). *Am. J. Epidemiol. 126,* 981–982.

Kushi, L. H., K. W. Samonds, J. M. Lacey, P. T. Brown, J. G. Bergan, and F. M. Sacks (1988). The association of dietary fat with serum cholesterol in vegetarians: The effect of dietary assessment on the correlation coefficient. *Am. J. Epidemiol. 128,* 1054–1064.

Kushi, L. H., R. A. Lew, F. J. Stare, C. R. Ellison, M. el Lozy, G. Bourke, L. Daly, I. Graham, M. Hickey, R. Malcahy, and J. Kevancy (1985). Diet and 20-year mortality from coronary heart disease: The Ireland-Boston diet-heart study. *N. Engl. J. Med. 312,* 811–818.

Lapidus, L., H. Anderson, C. Bengtsson, and I. Bosaeus (1986). Dietary habits in relation to incidence of cardiovascular disease and death in women: A 12-year follow-up of participants in the population study of women in Gothenburg, Sweden. *Am. J. Clin. Nutr. 44*, 444–448.

Lea, E. J., S. P. Jones, and D. V. Hamilton (1982). The fatty acids of erythrocytes of myocardial infarction patients. *Atherosclerosis 41*, 363–369.

Leaf, A. and P. C. Weber (1988). Cardiovascular effects of n-3 fatty acids. *N. Engl. J. Med. 318*, 549–557.

Levine, R. A., D. H. Streeten, and R. J. Doisy (1968). Effects of oral chromium supplementation on the glucose tolerance of elderly human subjects. *Metabolism 17*, 114–125.

Lewis, B. (1958). Composition of plasma cholesterol ester in relation to coronary-artery disease and dietary fat. *Lancet ii*, 71–73.

Lipid Research Clinics Program (1984). The lipid research clinics coronary primary prevention trial results. Reduction in incidence of coronary heart disease. *J. Am. Med. Assoc. 251*, 351–364.

Luoma, H., A. Aromaa, S. Helminen, H. Murtomaa, L. Kiviluota, S. Punsar, and P. Knekt (1983). Risk of myocardial infarction in Finnish men in relation to fluoride, magnesium, and calcium concentration in drinking water. *Acta Med. Scand. 213*, 171–176.

Mann, G. V. (1977). Diet-heart: end of an era. *N. Engl. J. Med. 297*, 644–650.

Manson, J. E., M. J. Stampfer, C. H. Hennekens, W. C. Willett (1987). Body weight and longevity. *J.A.M.A. 257*, 353–358.

Masironi, R., A. T. Miesch, M. D. Crawford, and E. I. Hamilton (1972). Geochemical environments, trace elements, and cardiovascular disease. *Bull. WHO 47*, 139–150.

McCarron, D. A., C. D. Morris, H. J. Henry, et al. (1984). Blood pressure and nutrient intake in the United States. *Science 224*, 1392–1398.

McGandy, R. B., D. M. Hegsted, and F. J. Stare (1967). Dietary fats, carbohydrates, and atherosclerotic vascular disease. *N. Engl. J. Med. 277*, 245–247.

Morrison, J. A., R. Larsen, L. Glatfelter, et al. (1980). Nutrient intake: relationships with lipids and lipoproteins in 6 19-year-old children: The Princeton School District Study. *Metabolism 29*, 133–140.

McGee, D. L., D. M. Reed, K. Yano, A. Kagan, and J. Tillotson (1984). Ten-year incidence of coronary heart disease in the Honolulu Heart Program: Relationship to nutrient intake. *Am. J. Epidemiol. 119*, 667–676.

McGill, H. C. (1968). *The geographic pathway of atherosclerosis*. Baltimore: Williams and Wilkins.

Meade, T. W., S. Mellows, M. Brozovic, et al. (1986). Haemostatic function and ischaemic heart disease: principal results of the Northwick Park Heart Study. *Lancet ii*, 533–537.

Meredith, A. P., P. E. Enterline, B. Peterson, and J. G. Pekover (1960). An epidemiologic diet study in North Dakota. *J. Am. Diet. Assoc. 37*, 339–343.

Miettinen, T. A., G. Alfthan, J. K. Huttunen, J. Pikkarainen, V. Naukkarinen, S. Mattila, and T. Kumlin (1983). Serum selenium concentration related to myocardial infarction and fatty acid content of serum lipids. *Brit. Med. J. 287*, 517–519.

Miettinen, T. A., V. Naukkarinen, J. K. Huttunen, S. Mattila, and T. Kumlin (1982). Fatty-acid composition of serum lipids predicts myocardial infarction. *Br. Med. J. 285*, 993–996.

Miller, N. E. (1987). Associations of high-density lipoprotein subclasses and apolipoproteins with ischemic heart disease and coronary atherosclerosis. *Am. Heart J. 113*, 589–597.

Moore, J. A., R. Novia, and I. C. Wells (1984). Selenium concentrations in plasma of patients with arteriographically defined coronary atherosclerosis. *Clin. Chem. 30,* 1171–1173.

Moore, R. D. and T. A. Pearson (1986). Moderate alcohol consumption and coronary artery disease: a review. *Medicine 65,* 242–267.

Morris, J. N., J. W. Marr, and D. G. Clayton (1977). Diet and heart: A postscript. *Br. Med. J. 2,* 1307–1314.

Multiple Risk Factor Interention Trial Research Group (1982). Multiple Risk Factor Intervention Trial. Risk Factor Changes and Mortality Results. *J. Am. Med. Assoc. 248,* 1465–1477.

Munro, H. N., M. H. Steele, and W. Forbes (1965). Effect of dietary protein level on deposition of cholesterol in the tissues of the cholesterol-fed rabbit. *Br. J. Exp. Pathol. 46,* 489–496.

National Diet Heart Study Research Group (1968). *National Diet Heart Study Final Report 37 (suppl 1),* 1–419.

Nichols, A. B., C. Ravenscroft, D. E. Lamphiear, and L. D. Ostrander Jr. (1976). Independence of serum lipid levels and dietary habits. The Tecumseh Study. *J.A.M.A. 236,* 1948–1953.

Norell, S. E., A. Ahlbom, M. Feychting, and N. L. Pedersen (1986). Fish consumption and mortality from coronary heart disease. *Brit. Med. J. 293,* 426.

Page, L. and R. M. Marston (1979). Food consumption pattern—U.S. Diet. In R. J. Havlik, M. Feinleib, eds.: *Proceedings of the conference on the decline in coronary heart disease mortality* U.S. Dept. Health, Education and Welfare, Public Health Service, NIH, Washington, D.C. (NIH Publ. No. 79-1610), pp. 236–243.

Palgi, A. (1981). Association between dietary changes and mortality rates: Israel 1949 to 1977; a trend-free regression model. *Am. J. Clin. Nutr. 34,* 1569–1583.

Potter, J. G., D. L. Topping, and D. Oakenfull (1979). Soya, saponins, and plasma cholesterol (letter). *Lancet i,* 223.

Potter, J. G., R. J. Illman, G. D. Calvert, D. G. Oakenfull, and D. L. Topping (1980). Soya, saponins, plasma lipids, lipoproteins, and fecal bile acids: A double-blind crossover study. *Nutrition Reports Int. 22,* 521–528.

Potter, J. G., K. P. Coffman, R. L. Reid, J. M. Krall, and J. M. Albrink (1981). Effect of test meals of varying dietary fiber content on plasma insulin and glucose response. *Am. J. Clin. Nutr. 34,* 328–334.

Ramirez, J. and N. C. Flowers (1980). Leukocyte ascorbic acid and its relationship to coronary artery disease in man. *Am. J. Clin. Nutr. 33,* 2079–2087.

Reiser, S., E. Bohn, J. Hallfrisch, O. E. Michaelis, 4th, M. Keeney, and E. S. Prather (1981). Serum insulin and glucose in hyperinsulinemic subjects fed three different levels of sucrose. *Am. J. Clin. Nutr. 34,* 2348–2358.

Renaud, S., F. Godsey, E. Dumont, C. Thevenon, E. Ortchanian, and J. L. Martin (1986). Influence of long-term diet modification on platelet function and composition in Moselle farmers. *Am. J. Clin. Nutr. 43,* 136–150.

Renaud, S., K. Kuba, C. Goulet, Y. Lemire, and C. Allard (1970). Relationship between fatty-acid composition of platelets and platelet aggregation in rat and man. Relation to thrombosis. *Circ. Res. 26,* 553–564.

Rentrop, K. P., F. Feit, H. Blanke, P. Stecy, R. Schneider, M. Rey, S. Horowitz, M. Goldman, K. Karsch, H. Meilman, et al. (1984). Effects of intracoronary streptokinase and intracoronary nitroglycerin infusion on coronary angiographic patterns and mortality in patients with acute myocardial infarction. *N. Engl. J. Med. 311,* 1457–1463.

Rhoads, G. G., C. L. Gulbrandsen, and A. Kagan (1976). Serum lipoproteins and coro-

nary heart disease in a population of Hawaiian Japanese men. *N. Engl. J. Med. 294,* 293–298.

Riales, R. and M. J. Albrink (1981). Effect of chromium chloride supplementation on glucose tolerance and serum lipids including high-density lipoprotein of adult men. *Am. J. Clin. Nutr. 34,* 2670–2678.

Riemersma, R. A., D. A. Wood, S. Butler, et al. (1986). Linoleic acid in adipose tissue and coronary heart disease. *Br. Med. J. 292,* 1423–1427.

Rinehart, J. F. and L. D. Greenberg (1951). Pathogenesis of experimental arteriosclerosis in pyidoxine deficiency. *Arch. Pathology 51,* 12–18.

Rinzler, S. H., H. Bakst, Z. H. Benjamin, A. L. Bobb, and J. Travell (1950). Failure of alpha tocopherol to influence chest pain in patients with heart disease. *Circulation 1,* 288–293.

Rivellese, A., G. Riccardi, A. Giacco, et al. (1980). Effect of dietary fibre on glucose control and serum lipoproteins in diabetic patients. *Lancet ii,* 447–450.

Robert, H. J. (1981). Perspective on vitamin E as therapy. *J. Am. Med. Assoc. 246,* 129–131.

Robertson, T. L., H. Kato, G. G. Rhoads, A. Kagan, M. Marmot, S. L. Syme, T. Gordon, R. M. Worth, J. L. Belsky, D. S. Dock, M. Miyanishi, and S. Kawamoto (1977). Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: Incidence of myocardial infarction and death from coronary heart disease. *Am. J. Cardiol. 39,* 239–243.

Rosenberg, L., D. Slone, S. Shapiro, D. W. Kaufman, O. S. Miettinen, and P. D. Stolley (1981). Alcoholic beverages and myocardial infarction in young women. *Am. J. Pub. Health 71,* 82–85.

Rowe, J. W., J. B. Young, K. L. Minaker, A. L. Stevens, J. Pallotta, and L. Landsberg (1981). Effect of insulin and glucose infusions on sympathetic nervous system activity in normal man. *Diabetes 30,* 219–225.

Sacks, F. M., J. L. Breslow, P. G. Wood, and E. H. Kass (1983). Lack of an effect of dairy protein (casein) and soy protein on plasma cholesterol of strict vegetarians. An experiment and a critical review. *J. Lipid Res. 24,* 1012–1020.

Salonen, J. T., G. Alfthan, J. K. Huttunen, J. Pikkarainen, and P. Puska (1982). Association between cardiovascular death and myocardial infarction and serum selenium in a matched-pair longitudinal study. *Lancet ii,* 175–179.

Schrade, W., R. Biegler, and E. Boehle (1961). Fatty-acid distribution in the lipid fractions of healthy persons of different age, patients with atherosclerosis and patients with idiopathic hyperlipidaemia. *J. Atheroscler. Res. 1,* 47–61.

Schroeder, H. A., A. P. Nason, and I. H. Tipton (1970). Chromium deficiency as a factor in atherosclerosis. *J. Chron. Dis. 23,* 123–142.

Scrimshaw, N. S. and M. A. Guzman (1968). Diet and Atherosclerosis. *Laboratory Investigation 18,* 623–628.

Seelig, M. (1980). *Magnesium Deficiency in the Pathogenesis of Disease.* New York: Plenum Publishing.

Seelig, M. and H. A. Heggtveit (1974). Magnesium interrelationships in ischemic heart disease: a review. *Am. J. Clin. Nutr. 27,* 59–79.

Shekelle, R. B., A. M. Shryock, O. Paul, et al. (1981). Diet, serum cholesterol and death from coronary heart disease. The Western Electric Study. *N. Engl. J. Med. 304,* 65–70.

Shekelle, R. B., M. Z. Nichaman, and W. J. Raynor (1987). Re: total energy intake: implication for epidemiologic analyses (letter). Am. J. Epidemiol. 126, 980.

Shekelle, R. B., L. Missell, O. Paul, A. M. Shyrock, and J. Stamler (1985). Fish consumption and mortality from coronary heart disease (letter). *N. Engl. J. Med. 313,* 820.

Simpson, H. C., K. Barker, R. D. Carter, E. Cassels, and J. I. Mann (1982). Low dietary intake of linoleic acid predisposes to myocardial infarction. *Br. Med. J. 285,* 683–684.

Sirtori, C. R., E. Gatti, O. Mantero, et al. (1979). Clinical experience with the soybean diet in the treatment of hypercholesterolemia. *Am. J. Clin. Nutr. 32,* 1645–1658.

Snowdon, D. A., R. L. Phillips, and G. E. Fraser (1984). Meat consumption and fatal ischemic heart disease. *Preventive Med. 13,* 490–500.

Stamler, J. (1985). The marked decline in coronary heart disease mortality rates in the United States, 1968–1981: Summary of findings and possible explanations. *Cardiology 72,* 11–22.

Stamler, J., D. Wentworth, and J. D. Neaton (1986). Is the relationship between serum cholesterol and risk of premature death from coronary heart disease continuous or graded? Findings in 356,222 primary screenees of the Multiple Risk Factor Intervention Trial (MRFIT). *J.A.M.A. 256,* 2823–2828.

Stampfer, M. J., G. A. Colditz, W. C. Willett, F. E. Speizer, and C. H. Hennekens (1988). A prospective study of moderate alcohol consumption and the risk of coronary heart disease and stroke in women. *N. Engl. J. Med. 319,* 267–273.

Steering Committee of the Physicians' Health Study Research Group (1988). Preliminary report: findings from the aspirin component of the ongoing Physicians' Health Study. *N. Engl. J. Med. 318,* 262–264.

Thomas, L. H. and R. G. Scott (1981). Ischaemic heart disease and the proportions of hydrogenated fat and ruminant-animal fat in adipose tissue at post-mortem examination: A case-control study. *J. Epidemiol. Commun. Health 35,* 251–255.

Turpeinen, O., M. J. Karronen, M. Pekkarinen, M. Miettinen, R. Elosuo, and E. Paavilainen (1979). Dietary prevention of coronary heart disease: The Finnish Mental Hospital Study. *Int. J. Epidemiol. 8,* 99–118.

Vahouny, G. V., W. Chalcarz, S. Satchithanandam, I. Adamson, D. M. Klurfeld, and D. Kritchevsky (1984). Effect of soy protein and casein intake on intestinal absorption and lymphatic transport of cholesterol and oleic acid. *Am. J. Clin. Nutr. 40,* 1156–1164.

Valimaki, M., E. A. Nikkila, M. R. Taskinen, and R. Ylikahri (1986). Rapid decrease in high density lipoprotein subfractions and postheparin plasma lipase activities after cessation of chronic alcohol intake. *Atherosclerosis 59,* 147–153.

Vermaak, W. J., H. C. Barnard, G. M. Potgieter, and J. D. Marx (1986). Plasma pyridoxal-5′-phosphate levels in myocardial infarction. *S. Afr. Med. J. 70,* 195–196.

Virtamo, J., E. Valkeila, G. Alfthan, S. Punsar, J. K. Huttunen, and M. J. Karvonen (1985). Serum selenium and the risk of coronary heart disease and stroke. *Am. J. Epidemiol. 122,* 276–282.

Vollset, S. E., I. Heuch, and E. Bjelke (1985). Fish consumption and mortality from coronary heart disease (letter). *N. Engl. J. Med. 313,* 820–821.

Wen, C. P. and S. N. Gershoff (1973). Changes in serum cholesterol and coronary heart disease mortality associated with changes in the post-war Japanese diet. *Am. J. Clin. Nutr. 26,* 616–619.

Willett, W. C. (1985). Does low vitamin B-6 intake increase the risk of coronary heart disease. In R. D. Reynolds, J. E. Leklem, eds.: *Vitamin B-6: Its Role in Health and Disease.* Current topics in nutrition and disease. Vol. 13. New York: Alan R. Liss, Inc., pp. 337–346.

Wissler, R. W. and D. Vesselinovitch (1975). The effects of feeding various dietary fats on the development and regression of hypercholesterolemia and atherosclerosis. *Adv. Exp. Med. Biol. 60,* 65–76.

Wood, D. A., R. A. Riemersma, S. Butler, M. Thomson, C. MacIntyre, R. A. Elton, and

M. F. Oliver (1987). Linoleic and eicosapentaenoic acids in adipose tissue and platelets and risk of coronary heart disease. *Lancet i,* 117–183.

Wood, D. A., S. Butler, R. A. Riemersma, M. Thomson, M. F. Oliver, M. Fulton, A. Birtwhistle, and R. Elton (1984). Adipose tissue and platelet fatty acids and coronary heart disease in Scottish men. *Lancet ii,* 117–121.

Working Group on Arteriosclerosis of the National Heart, Lung, and Blood Institute (1981). *Decline in coronary heart disease mortality, 1963–1978. Vol. 2.* Bethesda, Md: National Institutes of Health, 1981:157–258 (AHHS Publication No. (NIH) 82–2035).

Young, J. B. and L. Landsberg (1981). Effect of oral sucrose on blood pressure in the spontaneously hypertensive rat. *Metabolism 30,* 421–424.

Yudkin, J. (1963). Nutrition and palatability with special reference to obesity, myocardial infarction, and other diseases of civilization. *Lancet i,* 1335–1338.

16

Future Research Directions

In this brief chapter, an attempt is made to project the direction of nutritional epidemiology over the coming years and outline some of the research needed in this field. In some aspects, this direction is predictable due to studies or activities that are already in place. In other areas, the future is quite speculative and depends on scientific opportunities that are difficult to predict.

Probably the most important accomplishment of the last decade in nutritional epidemiology has been the development and validation of methods for measuring dietary intake that are sufficiently inexpensive to be used in large populations and yet accurate enough to provide informative answers to numerous existing hypotheses. Of fundamental importance is the fact that relatively simple structured questionnaires can discriminate among persons in a general population with respect to intake of a wide variety of nutrients; this means that their diets are not so homogeneous so as to preclude meaningful study. These developmental studies have set the stage for a new generation of prospective investigations during the next decade that are likely to produce a substantially greater body of reliable data than exists today.

DEVELOPMENT OF PROSPECTIVE COHORT DIETARY STUDIES

Prospective dietary studies are not a new phenomena; valuable studies, such as those of Hirayama (1979) and Shekelle and colleagues (1981), were started as long as 30 years ago. The older studies, however, have been limited by their narrow scope of questions (for example, Hirayama's questionnaire included only five food items), the use of 24-hour recalls to assess dietary intake, study populations that were too small to assess any but the most common conditions, or lack of reassessment of diet during prolonged follow-up periods. A newer generation of prospective dietary studies now exists that are based on self-admin-

istered food-frequency questionnaires completed by groups sufficiently large to allow important cancer sites as well as cardiovascular outcomes to be examined. In North America, at least six large cohort studies with reasonably complete dietary assessments have been established: the California Seventh Day Adventists study of approximately 40,000 men and women (Gary Frazer, personal communication), the Nurses' Health Study of approximately 95,000 women, the New York State cohort of approximately 80,000 men and women (Saxon Graham, personal communication), the Canadian breast cancer screening study of approximately 56,000 women (Jain et al., 1982), the Health Professionals Follow-up Study of approximately 52,000 men (which is conducted in parallel with the Nurses' Health Study), and the Iowa Womens' Health Study of approximately 42,000 women (Aaron Folsom, personal communication). Because three of these cohorts were designed to study specific cancers of women, men in these studies are substantially outnumbered by women, in contrast to most of the existing data that are derived from studies primarily designed to investigate coronary heart disease in men.

Large cohort studies of diet have been started or are being planned in other countries as well. The prospective study begun by Bjelke in Norway, although not as large or comprehensive as the newer cohorts, has already provided important information of vitamin A intake and lung cancer (Kvale et al., 1983). In Holland, approximately 130,000 men and women have been enrolled in a prospective study (Piet van den Brandt, personal communication). Studies are in various stages of planning or implementation in Denmark, Sweden, Japan, France, and the Soviet Union, and consideration is being given to the development of dietary cohorts in Australia and Germany as well.

Collectively, the ongoing and planned prospective studies provide a wealth of new data on diet in relation to cancer, coronary heart disease, and other outcomes; these data are far less subject to the methodologic biases that can affect case-control studies, and should provide a more coherent and consistent picture of these relationships than is available today. As overall diet and disease relationships become well established, these studies should also provide relatively precise estimates of dose-response relationships and interactions with other dietary and nondietary factors. The future body of literature could be enhanced by additional cohort studies among populations with diets other than those of typical Western cultures, and by periodic reassessments of diet during the course of follow-up.

VALIDATION STUDIES

Validation studies, in which a dietary questionnaire is compared with a detailed assessment method such as diet recording or with biochemical markers, are being conducted within most of the recently begun prospective cohort investigations. These substudies, which could also be called standardization or calibration studies, should be a component of any new major dietary study. In future reports from prospective studies, the validation substudy can provide a better estimate of the actual distributions of nutrient intakes within the study popu-

lation. If only the data from a questionnaire are used, the true variation among subjects cannot be separated from variation due to measurement error.

If no association is seen in a particular study, the validation study permits a distinction between the possibilities that this is due to lack of variation in the dietary factor, that the questionnaire is insufficiently accurate, or that no association really exists within defined constraints of time and variation in diet. Although almost no epidemiologic studies have presented relative risks and confidence intervals corrected for measurement error, use of data from validation studies and relatively simple statistical methods for correcting these measures of association encourage the routine presentation of corrected results. Validation studies also allow the findings from multiple investigations to be combined in a manner that reflects the accuracy of the dietary data from each of the individual studies.

Many aspects of dietary questionnaire validity deserve further investigation. The performance of these questionnaires in other populations is of interest, particularly among less educated groups and the populations of developing countries. An extention of validation studies to additional nutrients and nonnutritive aspects of diet, such as food additives and the chemical products of cooking and processing food, will also be useful. The use of additional biochemical markers for standards is likely to be particularly helpful for some nutrients for which the assumptions involved in computing intakes are open to question. For example, even though folic acid calculated from questionnaires correlates reasonably with calculated intake from diet records, both intakes may be erroneous due to the lability of this nutrient; an assessment of how either method predicts blood levels would greatly enhance our interpretation of calculated intakes.

CASE-CONTROL STUDIES IN NUTRITIONAL EPIDEMIOLOGY

The future role of case-control studies in epidemiologic studies of diet remains uncertain. For some relationships, such as that for green and yellow vegetable intake in relation to lung cancer, the results of studies have been remarkably consistent, but for other relationships clear answers have not emerged. Perhaps many or most of the case-control studies have provided valid, unbiased estimates of dietary relationships in their study populations. The potential for biases in the selection of controls and in the recall of past diet, however, makes it difficult to be confident in the results of any one study or even in the pooled results of many studies. When results are conflicting, some studies are presumably correct and some incorrect; unfortunately, it is usually difficult to determine which is which.

Several studies are now ongoing in which biases in the recall of past diet are being assessed by comparing responses of cases and controls with previously collected dietary data. These studies provide useful information for the interpretation of case-control studies. The problems of selecting appropriate controls, however, and the low participation rates that have plagued many recent studies still need to be addressed, as willingness to provide information may well be related to dietary intake. Because many less common diseases cannot be studied

adequately even in the larger cohorts, case-control studies may provide the only opportunity to investigate some dietary relationships. It is, therefore, essential to define better the limitations of this methodology.

BIOCHEMICAL MARKERS OF DIET

Studies of diet and disease could be substantially enhanced by new and better biochemical markers of intake. The use and interpretation of many existing markers can be improved by a better characterization of their relationship with dietary intake and their variance components between persons and within persons over different periods of time. The search for new markers of diet should concentrate on those that are likely to provide time-integrated measures, such as protein adducts and levels in fat stores or nails.

Despite general enthusiasm for the use of biochemical markers of diet, they are not likely to replace dietary questionnaires in nutritional epidemiology as practical markers are not in sight for many nutrients of greatest interest and as dietary data can be collected for large populations at far less cost using self-administered questionnaires. For the latter reason, biochemical markers may play a greater role in calibrating and improving questionnaires than in direct application in case-control or cohort studies.

USE OF INTERMEDIATE AND PREMORBID ENDPOINTS

Future studies of dietary effects are likely to use premorbid endpoints more often as many of the methodologic problems inherent in traditional case-control studies may be reduced or eliminated. Examples of these endpoints could include dysplastic lesions, in situ neoplasms, or premalignant growths such as colon polyps. In the area of cardiovascular disease, epidemiologists have frequently studied determinants of risk factors for coronary heart disease, such as hypertension. More recently, noninvasive methods are providing measurements of atherosclerosis at different sites in the body. For these endpoints, it is much less likely that the disease or symptoms will affect diet or biochemical markers of diet. Moreover, studies can be designed so that subjects are not aware of whether they are cases or controls.

The application of expensive or unstable biochemical markers of exposure may be particularly well suited to case-control studies with premorbid endpoints as opposed to large cohort studies with clinical disease as an endpoint. In addition, intervention trials are making increased use of these endpoints as the number of subjects and follow-up time required are usually far less than for trials to prevent clinical outcomes. Nevertheless, studies of premorbid endpoints cannot completely substitute for investigations with the event of ultimate interest as the outcome because the effect on clinical disease may not always be predicted by the effect on a benign or premorbid lesion. For example, oral contraceptives have been shown to reduce the occurrence of benign breast lesions in many studies, but do not appear to reduce the risk of breast cancer. For some diseases

concordant findings from a combination of observational data with clinical disease as the outcome, as might be obtained from cohort studies, and intervention studies with premalignant lesions as an endpoint would be extremely compelling evidence of a causal relationship.

Recent advances in molecular and cell biology are providing new possibilities for future studies of dietary effects. A variety of measurements, such as excision products of DNA measured in the urine (Cathcart et al., 1984), sister chromatid exchange, and micronuclei counts, provide indicators of presumed damage to DNA (Rosen et al., 1987). These may be used as outcomes to evaluate the effect of factors that are presumed either to reduce or increase damage to macromolecules. The identification of specific growth factors for malignant cells may provide additional short-term indicators for the effects of diet. The expression of specific oncogenes would conceptually provide an interesting endpoint for study, but as they are usually identified only in tumors themselves, their epidemiologic use may be limited to creating subcategories of malignancies rather than as premorbid endpoints. Although interesting research opportunities are likely to exist at the interface of nutritional epidemiology and basic biology during the coming years, the contribution that these will make to our knowledge of diet and disease remains unclear at this time.

ROLE OF NUTRITIONAL EPIDEMIOLOGY IN TRIALS

As dietary hypotheses become better defined and supported, an increasing number are likely to be evaluated by intervention trials. Nutritional epidemiology will play an important role in these trials. One important role will be to screen or identify persons for eligibility. As discussed earlier, the ability of a study to detect an effect of diet will be greatest among those who have not yet partially adopted the intervention; that is, those with high intake in a trial to reduce a dietary factor and those with low intake in a trial to increase a dietary factor. Therefore, simple assessment methods to identify those with diets most susceptible to intervention should be incorporated in such future trials.

Monitoring dietary compliance, both at the individual and treatment group levels, presents great challenges as persons who are being taught to modify their diet may well report better compliance than is true. If dietary records are used, adherence to the diet may be better while keeping records than at other times. Because the possibility for biased information is especially great in this situation, objective markers of diet, such as biochemical measurements, will be especially useful.

INTERACTION WITH GENETIC FACTORS

Although the occurrence of many diseases, such as breast and colon cancers and coronary heart disease, represents an interaction between genetic and environmental factors, such as diet, these interactions have seldom been investigated in case-control or cohort studies. A major limitation has been that the genetic pre-

disposition has not been directly measurable; usually only family history of the disease has been available. Rapid progress is currently being made in the identification of specific genes for susceptibility to many important diseases, and in the development of DNA probes to classify individuals according to the presence or absence of these predisposing genes. The availability of such probes, which usually require only white blood cells for examination (Gustafson et al., 1987), can add powerfully to studies of diet and disease. If, for example, an effect of diet is limited to a subset of genetically susceptible persons, the relative risks in this subset should be substantially higher than in the population as a whole. Furthermore, studies with information on both dietary intake and genetic predisposition may be able to identify large segments of the population for whom specific dietary changes are not important. When such studies become available, they will add an important and exciting dimension to our knowledge.

SUMMARY

The new generation of large cohort studies of diet that have already been started or are being planned will provide a wealth of additional, high-quality information on the relationships of diet with common diseases. Validation studies embedded in most of these new studies will allow quantitive estimates of associations unattenuated by measurement error. The use of premorbid endpoints and newer markers of tissue damage can enhance the plausibility of findings from studies with clinical disease as an endpoint and improve our understanding of the pathophysiology involved. The identification of individuals at high genetic risk of disease using DNA probes can create new opportunities to study interactions between genetic factors and diet.

REFERENCES

Cathcart, R., E. Schwiers, R. L. Saul, and B. N. Ames (1984). Thymine glycol and thymidine glycol in human and rat urine; A possible assay for oxidative DNA damage. *Proc. Natl Acad. Sci. USA 81,* 5633–5637.

Gustafson, S., J. A. Proper, E. J. Bowie, and S. S. Sommer (1987). Parameters affecting the yield of DNA from human blood. *Anal. Chem. 165,* 294–299.

Hirayama, T. (1979). Diet and cancer. *Nutr. Cancer 1,* 67–81.

Jain, M. G., L. Harrison, G. R. Howe, and A. B. Miller (1982). Evaluation of a self-administered dietary questionnaire for use in a cohort study. *Am. J. Clin. Nutr. 36,* 931–935.

Kvale, G., E. Bjelke, and J. J. Gart (1983). Dietary habits and lung cancer risk. *Int. J. Cancer 31,* 397–405.

Rosin, M. P., B. P. Dunn, and H. F. Strich (1987). Use of intermediate endpoints in quantitating the response of precancerous lesions to chemopreventive agents. *Can. J. Physiol. Pharmocol. 65,* 483–487.

Shekelle, R. B., A. M. Shryock, O. Paul, et al. (1981). Diet, serum cholesterol, and death from coronary heart disease. The Western Electric Study. *N. Engl. J. Med 304,* 65–70.

Index

ISBN 0-19-504501-7