Biostatistical Methods

The Assessment of Relative Risks

JOHN M. LACHIN
The George Washington University
Washington, D.C.

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

This Page Intentionally Left Blank

Biostatistical Methods

WILEY SERIES IN PROBABILITY AND STATISTICS
APPLIED PROBABILITY AND STATISTICS SECTION

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *Vic Barnett, Noel A. C. Cressie, Nicholas I. Fisher, Iain M. Johnstone, J. B. Kadane, David W. Scott, Bernard W. Silverman, Adrian F. M. Smith, Jozef L. Teugels; Ralph A. Bradley, Emeritus, J. Stuart Hunter, Emeritus, David G. Kendall, Emeritus*

A complete list of the titles in this series appears at the end of this volume.

Biostatistical Methods

The Assessment of Relative Risks

JOHN M. LACHIN
The George Washington University
Washington, D.C.

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

To my family

This text is printed on acid-free paper. ♾

Published simultaneously in Canada.

To order books or for customer service please, call 1(800)-CALL-WILEY (225-5945).

Library of Congress Cataloging-in-Publication Data is available.

ISBN 0-471-36996-9

Printed in the United States of America

10 9 8 7 6 5 4 3 2

Contents

Preface

In 1993 to 1994 I led the effort to establish a graduate program in biostatistics at the George Washington University. The program, which I now direct, was launched in 1995 and is a joint initiative of the Department of Statistics, the Biostatistics Center (which I have directed since 1988) and the School of Public Health and Health Services. Biostatistics has long been a specialty of the statistics faculty, starting with Samuel Greenhouse, who joined the faculty in 1946. When Jerome Cornfield joined the faculty in 1972, he established a two-semester sequence in biostatistics (Statistics 225-6) as an elective for the graduate program in statistics (our 200 level being equivalent to the 600 level in other schools). Over the years these courses were taught by many faculty as a lecture course on current topics. With the establishment of the graduate program in biostatistics, however, these became pivotal courses in the graduate program and it was necessary that Statistics 225 be structured so as to provide students with a review of the foundations of biostatistics.

Thus I was faced with the question "what are the foundations of biostatistics?" In my opinion, biostatistics is set apart from other statistics specialties by its focus on the assessment of risks and relative risks through clinical research. Thus biostatistical methods are grounded in the analysis of binary and count data such as in 2×2 tables. For example, the Mantel-Haenszel procedure for stratified 2×2 tables forms the basis for many families of statistical procedures such as the G^ρ family of modern statistical tests in the analysis of survival data. Further, all common medical study designs, such as the randomized clinical trial and the retrospective case-control study, are rooted in the desire to assess relative risks. Thus I developed

Statistics 225, and later this text, around the principle of the assessment of relative risks in clinical investigations.

In doing so, I felt that it was important first to develop basic concepts and derive core biostatistical methods through the application of classical mathematical statistical tools, and then to show that these and comparable methods may also be developed through the application of more modern, likelihood-based theories. For example, the large sample distribution of the Mantel-Haenszel test can be derived using the large sample approximation to the hypergeometric and the Central Limit Theorem, and also as an efficient score test based on a hypergeometric likelihood.

Thus the first five chapters present methods for the analysis of single and multiple 2×2 tables for cross-sectional, prospective and retrospective (case-control) sampling, without and with matching. Both fixed and random effects (two-stage) models are employed. Then, starting in Chapter 6 and proceeding through Chapter 9, a more modern likelihood or model-based treatment is presented. These chapters broaden the scope of the book to include the unconditional and conditional logistic regression models in Chapter 7, the analysis of count data and the Poisson regression model in Chapter 8, and the analysis of event time data including the proportional hazards and multiplicative intensity models in Chapter 9. Core mathematical statistical tools employed in the text are presented in the Appendix. Following each chapter problems are presented that are intended to expose the student to the key mathematical statistical derivations of the methods presented in that chapter, and to illustrate their application and interpretation.

Although the text provides a valuable reference to the principal literature, it is not intended to be exhaustive. For this purpose, readers are referred to any of the excellent existing texts on the analysis of categorical data, generalized linear models and survival analysis. Rather, this manuscript was prepared as a textbook for advanced courses in biostatistics. Thus the course (and book) material was selected on the basis of its current importance in biostatistical practice and its relevance to current methodological research and more advanced methods. For example, Cornfield's *approximate* procedure for confidence limits on the odds ratio, though brilliant, is no longer employed because we now have the ability to readily perform exact computations. Also, I felt it was more important that students be exposed to over-dispersion and the use of the information sandwich in model-based inference than to residual analysis in regression models. Thus each chapter must be viewed as one professor's selection of relevant and insightful topics.

In my Statistics 225 course, I cover perhaps two-thirds of the material in this text. Chapter 9, on survival analysis, has been added for completeness, as has the section in the Appendix on quasi-likelihood and the family of generalized linear models. These topics are covered in detail in other courses. My detailed syllabus for Statistics 225, listing the specific sections covered and exercises assigned, is available at the Biostatistics Center web site (www.bsc.gwu.edu/jml/biostatmethods). Also, the data sets employed in the text and problems are available at this site or the web site of John Wiley and Sons, Inc. (www.wiley.com).

Although I was not trained as a mathematical statistician, during my career I have learned much from those with whom I have been blessed with the opportunity

to collaborate (chronologically): Jerry Cornfield, Sam Greenhouse, Nathan Mantel, and Max Halperin, among the founding giants in biostatistics; and also Robert Smythe, L.J. Wei, Peter Thall, K.K. Gordon Lan and Zhaohai Li, among others, who are among the best of their generation. I have also learned much from my students, who have always sought to better understand the rationale for biostatistical methods and their application.

I especially acknowledge the collaboration of Zhaohai Li, who graciously agreed to teach Statistics 225 during the fall of 1998, while I was on sabbatical leave. His detailed reading of the draft of this text identified many areas of ambiguity and greatly improved the mathematical treatment. I also thank Costas Cristophi for typing my lecture notes, and Yvonne Sparling for a careful review of the final text and programming assistance. I also wish to thank my present and former statistical collaborators at the Biostatistics Center, who together have shared a common devotion to the pursuit of good science: Raymond Bain, Oliver Bautista, Patricia Cleary, Mary Foulkes, Sarah Fowler, Tavia Gordon, Shuping Lan, James Rochon, William Rosenberger, Larry Shaw, Elizabeth Thom, Desmond Thompson, Dante Verme, Joel Verter, Elizabeth Wright, and Naji Younes, among many.

Finally, I especially wish to thank the many scientists with whom I have had the opportunity to collaborate in the conduct of medical research over the past 30 years: Dr. Joseph Schachter, who directed the Research Center in Child Psychiatry where I worked during graduate training; Dr. Leslie Schoenfield, who directed the National Cooperative Gallstone Study; Dr. Edmund Lewis, who directed the Collaborative Study Group in the conduct of the Study of Plasmapheresis in Lupus Nephritis and the Study of Captropil in Diabetic Nephropathy; Dr. Thomas Garvey, who directed the preparation of the New Drug Application for treatment of gallstones with ursodiol; Dr. Peter Stacpoole, who directed the Study of Dichloroacetate in the Treatment of Lactic Acidosis; and especially Drs. Oscar Crofford, Saul Genuth and David Nathan, among many others, with whom I have collaborated since 1982 in the conduct of the Diabetes Control and Complications Trial, the study of the Epidemiology of Diabetes Interventions and Complications, and the Diabetes Prevention Program. The statistical responsibility for studies of such great import has provided the dominant motivation for me to continually improve my skills as a biostatistician.

JOHN M. LACHIN

Rockville, Maryland

This Page Intentionally Left Blank

1

Biostatistics and Biomedical Science

1.1 STATISTICS AND THE SCIENTIFIC METHOD

The aim of all biomedical research is the acquisition of new information so as to expand the body of knowledge that comprises the biomedical sciences. This body of knowledge consists of three broad components:

1. Descriptions of phenomena in terms of observable characteristics of elements or events;

2. Descriptions of associations among phenomena;

3. Descriptions of causal relationships between phenomena.

The various sciences can be distinguished by the degrees to which each contains knowledge of each of these three types. The hard sciences (e.g. physics and chemistry) contain large bodies of knowledge of the third kind — causal relationships. The soft sciences (e.g. the social sciences) principally contain large bodies of information of the first and second kind — phenomenological and associative.

None of these descriptions, however, are exact. To quote the philosopher and mathematician Jacob Bronowski (1973).

> All information is imperfect. We have to treat it with humility... Errors are inextricably bound up with the nature of human knowledge...

Thus every science consists of shared information, all of which to some extent is uncertain.

When a scientific investigator adds to the body of scientific knowledge, the degree of uncertainty about each piece of information is described through statistical assessments of the probability that statements are either true or false. Thus the language of science is statistics, for it is through the process of statistical analysis and interpretation that the investigator communicates the results to the scientific community. The syntax of this language is probability, because the laws of probability are used to assess the inherent uncertainty, errors, or precision of estimates of population parameters, and probabilistic statements are used as the basis for drawing conclusions.

The means by which the investigator attempts to control the degree of uncertainty in the research conclusions is the application of the scientific method. In a nutshell, the scientific method is a set of strategies, based on common sense and statistics, that is intended to minimize the degree of uncertainty and maximize the degree of validity of the resulting knowledge. Therefore, the scientific method is deeply rooted in statistical principles.

When considered sound and likely to be free of error, such knowledge is termed scientifically valid. The designation of scientific validity, however, is purely subjective. The soundness or validity of any scientific result depends on the manner in which the observations were collected, that is, on the design and conduct of the study, as well as the manner in which the data were analyzed.

Therefore, in the effort to acquire scientifically valid information, one must consider the statistical aspects of all elements of a study – its design, execution and analysis. To do so requires a firm understanding of the statistical basis for each type of study and for the analytic strategies commonly employed to assess a study's objectives.

1.2 BIOSTATISTICS

Biostatistics is principally characterized by the application of statistical principles to the biological/biomedical sciences; in contrast to other areas of application of statistics, such as psychometrics and econometrics. Thus biostatistics refers to the development of statistical methods for, and the application of statistical principles to, the study of biologic and medical phenomena.

Biomedical research activities range from the study of cellular biology to clinical therapeutics. At the basic physical level it includes so-called bench research or the study of genetic, biochemical, physiologic, and biologic processes, such as the study of genetic defects, metabolic pathways, kinetic models and pharmacology. Although some studies in this realm involve investigation in animals and man (*in vivo*), many of these investigations are conducted in "test tubes" (*in vitro*). The ultimate objective of these inquiries is to advance our understanding of the pathobiology or pathophysiology of diseases in man and of the potential mechanisms for their treatment.

Clinical research refers to the direct observation of the clinical features of populations. This includes *epidemiology*, which can be broadly defined as the study

of the distribution and etiology of human disease. Some elements, such as infectious disease epidemiology, are strongly biologically based, whereas others are more heavily dependent on empirical observations within populations. These latter include such areas as occupational and environmental epidemiology or the study of the associations between occupational and environmental exposures with the risk of specific diseases. This type of epidemiology is often characterized as *population-based* because it relies on the observation of natural samples from populations.

Ultimately, bench research or epidemiologic observation leads to advances in medical therapeutics — the development of new pharmaceuticals (drugs), devices, surgical procedures or interventions. Such therapeutic advances are often assessed using a randomized, controlled, clinical trial. Such studies evaluate the biological effectiveness of the new agent (biological efficacy), the clinical effectiveness of the therapy in practice (the so-called intention-to-treat comparison), as well as the incidence of adverse effects.

The single feature that most sharply distinguishes clinical biomedical research from other forms of biological research is the propensity to assess the absolute and relative risks of various outcomes within populations. The *absolute risk* refers to the distribution of a disease, or risk factors for a disease, in a population. This risk may be expressed cross-sectionally as a simple probability, or it may be expressed longitudinally over time as a hazard function (or survival function) or an intensity process. The *relative risk* refers to a measure of the difference in risks among subsets of the population with specific characteristics, such as those exposed versus not to a risk factor, or those randomly assigned to a new drug treatment versus a placebo control. The relative risk of an outcome is sometimes described as a difference in the absolute risks of the outcome, the ratio of the risks, or a ratio of the odds of the outcome.

Thus a major part of biostatistics concerns the assessment of absolute and relative risks through epidemiologic studies of various types and randomized clinical trials. This, in general, is the subject of this text. This entails the study of discrete outcomes, some of which are assessed over time. This also includes many major areas of statistics that are beyond the scope of any single text. For example, the analysis of longitudinal data is another of the various types of processes studied through biostatistics. In many studies, however, interest in a longitudinal quantitative or ordinal measure arises because of its fundamental relationship to an ultimate discrete outcome of interest. For example, longitudinal analysis of serum cholesterol levels in a population is of interest because of the strong relationship between serum lipids and the risk of cardiovascular disease, not cholesterol itself. Thus this text is devoted exclusively to the assessment of the risks of discrete characteristics or events in populations.

1.3 NATURAL HISTORY OF DISEASE PROGRESSION

Underlying virtually all clinical research is some model of our understanding of the natural history of the progression of the disease under investigation. As an example,

Table 1.1 Stages of Progression of Diabetic Nephropathy

1. Normal: Albumin excretion rate (AER) $\leq$ 40 mg/24 h
2. Microalbuminuria: 40 < AER < 300 mg/24 h
3. Proteinuria (overt albuminuria): AER $\geq$ 300 mg/24 h
4. Renal insufficiency: Serum creatinine > 2 mg/dL
5. End-stage renal disease: Need for dialysis or renal transplant
6. Mortality

consider the study of diabetic nephropathy (kidney disease) associated with type 1 or insulin dependent diabetes mellitus (IDDM), also known as juvenile diabetes. Diabetes is characterized by a state of metabolic dysfunction in which the subject is deficient in endogenous (self-produced) insulin. Thus the patient must administer exogenous insulin by some imperfect mechanical device, such as by multiple daily injections or a continuous subcutaneous insulin infusion (CSII) device also called a "pump". Because of technological deficiencies with the way insulin can be administered, it is difficult to maintain normal levels of blood glucose throughout the day, day after day. The resulting hyperglycemia leads to microvascular complications, the two most prevalent being diabetic retinopathy (disease of the retina in the eye) and diabetic nephropathy, and ultimately to cardiovascular disease.

Diabetic nephropathy is known to progress through a well-characterized sequence of disease states, characterized in Table 1.1. The earliest sign of emergent kidney disease is the leakage of small amounts of protein (albumin) into urine. The amount or rate of albumin excretion can be measured from a timed urine collection in which all the urine voided over a fixed period of time is collected. From the measurement of the urine volume and the concentration of albumin in the serum and urine at specific intervals of time, it is possible to compute the albumin excretion rate (AER) expressed as the mg/24 h of albumin excreted into the urine by the kidneys.

In the normal (non-diseased) subject, the AER is no greater than 40 mg/24 h, some would say no greater than 20 or 30 mg/24 h. The earliest sign of possible diabetic nephropathy is microalbuminuria, defined as an AER >40 mg/24 h (but < 300 mg/24 h). As the disease progresses, the next landmark is the development of definite albuminuria, defined as an AER >300 mg/24 h. This is often termed overt proteinuria because it is at this level of albumin (protein) excretion that a simple dip-stick test for protein in urine will be positive. This is also the point at which nephropathy, and the biological processes that ultimately lead to destruction of the kidney, are considered well established.

To then chart the further loss of kidney function, a different measure is used — the glomerular filtration rate (GFR). The glomerulus is the cellular structure that serves as the body's filtration system. As diabetic nephropathy progresses, fewer and fewer intact glomeruli remain, so that the rate of filtration declines, starting with the leakage of protein and other elements into the urine. The GFR is difficult to measure accurately. In practice, a measure of creatinine clearance, also from a timed urine collection, or a simple measure of the creatinine concentration in serum are used to monitor disease progression. Renal insufficiency is often declared when

the serum creatinine exceeds 2 mg/dL. This is followed by end-stage renal disease (ESRD), at which point the patient requires frequent dialysis or renal transplantation to prolong survival. Ultimately the patient dies from the renal insufficiency or related causes if a suitable donor kidney is not available for transplantation.

Thus the natural history of diabetic nephropathy is described by a collection of quantitative, ordinal and qualitative assessments. In the early stages of the disease, a study might focus entirely on quantitative measures of AER. Later, during the middle stages of the disease, this becomes problematic. For example, patients with established proteinuria may be characterized over time using a measure of GFR, but the analysis will be complicated by informatively missing observations because some patients reached ESRD or died before the scheduled completion of follow-up.

However, a study that assesses the risk of discrete outcomes, such as the incidence or prevalence of proteinuria or renal insufficiency, is less complicated by such factors and is readily interpretable by physicians. For example, if a study shows that a new drug treatment reduces the mean AER by 10 mg/24 h less than that with placebo, it is difficult to establish the clinical significance of the result. On the other hand, if the same study demonstrated a relative risk of developing proteinuria of 0.65, a 35% risk reduction with drug treatment versus placebo, the clinical significance is readily apparent to most physicians.

Therefore, we shall focus on the description of the absolute and relative risks of discrete outcomes, historically the core of biostatistics.

1.4 TYPES OF BIOMEDICAL STUDIES

Biomedical research employs various types of study designs, some of which involve formal experimentation, others not, among other characteristics. In this section the characteristics and the roles of each type of study are briefly described.

Study designs can be distinguished by three principal characteristics:

1. *Number of samples*: single versus multiple samples;

2. *Source of samples*: natural versus experimental. An experimental sample is one to which a treatment or procedure has been applied by the investigator. This may or may not involve randomization as an experimental device to assign treatments to individual patients.

3. *Time course of observation*: prospective versus retrospective versus concurrent collection of measurements and observation of responses or outcome events.

Based on these characteristics, there are basically four types of designs for biomedical studies in man: (1) the cross-sectional study, (2) the cohort study, (3) the case-control study, and (4) the randomized experiment. A more exhaustive classification was provided by Bailar, Louis, Lavori and Polansky (1984), but these four are the principal types. Examples of each type of study are described subsequently.

The **cross-sectional study** is a study of a single, natural sample with concurrent measurement of a variety of characteristics. In the review by Bailar, Louis, Lavori, and Polansky (1984), 39% of published studies were of this type. Some notable examples are the National Health and Nutritional Examination Survey (NHANES) of the relationship between health and nutrition, and the annual Health Interview Survey of the prevalence of various diseases in the general U.S. population. Such studies have provided important descriptions of the prevalence of disease in specified populations, of the co-occurrence of the disease and other factors (i.e. associations), and of the sensitivity and specificity of diagnostic procedures.

In a **cohort study** (25% of studies), one or more samples (cohorts) of individuals, either natural or experimental samples, are followed prospectively and subsequent status is evaluated.

A **case-control study** (5% of studies) employs multiple, natural samples with retrospective measurements. A sample of cases with the disease is compared to a sample of controls without the disease with respect to the previous presence of, or exposure to, some factor.

An important characteristic of cohort and case-control studies is whether or not the study employs **matching** of pairs or sets of subjects with respect to selected covariate values. Matching is a strategy to remove bias in the comparison of groups by ensuring equality of distributions of the selected matching covariates. Matching, however, changes the sample frame or the sampling unit in the analysis from the individual subject in an unmatched study to the matched set in the matched study. Thus matched studies require analytic procedures that are different from those more commonly applied to unmatched studies.

A **randomized, controlled clinical trial or parallel – comparative trial** (15% of studies) employs two or more parallel randomized cohorts, each of which receives only one treatment in the trial. Such studies provide a controlled assessment of a new drug, therapy, diagnostic procedure, or intervention procedure. Variations of this design include the multiple-period **crossover** design and the crossed **factorial** design. Since a clinical trial uses randomization to assign each subject to receive either the active treatment versus a control (e.g. drug vs. placebo), the comparison of the groups is in expectation unbiased. However, a truly unbiased study also requires other conditions such as complete and unbiased follow-up assessments.

Each of the first three types are commonly referred to as an observational or epidemiological study, in contrast to the clinical trial. It is rare, some might say impossible, that a population-based observational study will identify a single necessary and sufficient cause for a biologic effect, or a 1:1 causal relationship. Almost always, a *risk factor* is identified that has a biological effect that is associated with a change in the risk of an outcome. It is only after a preponderance of evidence is accumulated from many such studies that such a risk factor may be declared to be a *causal agent.* Such was the case with the relationship between smoking and lung cancer, and the criteria employed to declare smoking a causal agent are now widely accepted (US Surgeon General, 1964, 1982).

The principal advantage of the randomized controlled trial (RCT), on the other hand, is that it can provide conclusions with respect to causal relationships because

other intervening factors are controlled through randomization. Thus the RCT provides an unbiased comparison of the effects of administering one treatment versus another on the outcome in the selected population of patients, and any differences observed can be confidently ascribed to the differences between the treatments. Therefore, the distinction between a relationship based on an observational study and one based on a randomized experiment rests in the degree to which an observed relationship might be explained by other variables or other mechanisms.

However, in no study is there an absolute guarantee that all possible influential variables are controlled, even in a randomized, controlled experiment. Also, as the extent of knowledge about the underlying natural history of a disease expands, it becomes increasingly important to account for the known or suspected risk factors in the assessment of the effects of treatments or exposures, especially in an observational cross-sectional, cohort, or case-control study. This entails the use of an appropriate statistical model for the simultaneous influence of multiple covariates on the absolute or relative risk of important outcomes or events.

Thus the principal objective of this text is to describe methods for the assessment of risk relationships derived from each type of study, and to consider methods to adjust or control for other factors in these assessments.

1.5 STUDIES OF DIABETIC NEPHROPATHY

To illustrate the different types of studies, we close this chapter with a review of selected studies on various aspects of diabetic nephropathy.

Cross-sectional surveys such as the National Health Interview Survey (NHIS) and the National Health and Nutrition Evaluation Survey (NHANES) indicate that approximately 16 million people in the United States population have some form of diabetes mellitus (Harris, Hadden, Knowler and Bennett, 1987). The majority have what is termed type 2 or non-insulin dependent diabetes mellitus (NIDDM). Approximately 10% or 1.6 million have the more severe form termed type 1 or insulin-dependent diabetes mellitus (IDDM) for which daily insulin injections or infusions are required to sustain life. Among the most important clinical features of type 1 diabetes are the development of complications related to micro- and macrovascular abnormalities, among the most severe being diabetic nephropathy (kidney disease), which ultimately leads to end-stage renal disease (ESRD) in about a third of patients. These and other national surveys indicate that approximately 35% of all ESRD in the United States is attributed to diabetes.

As an illustration of a longitudinal observational cohort study, Deckert et al. (1978) followed a cohort of 907 Danish subjects with type 1 diabetes for many years and reported the annual incidence (proportion) of new cases of proteinuria (overt albuminuria) to appear each year. They showed that the peak incidence or greatest risk occurs approximately 15 years after the onset of diabetes. Their study also showed that over a lifetime, approximately 70% of subjects develop nephropathy whereas approximately 30% do not, suggesting that there is some mechanism that protects patients from nephropathy, possibly of a genetic nature, possibly related to

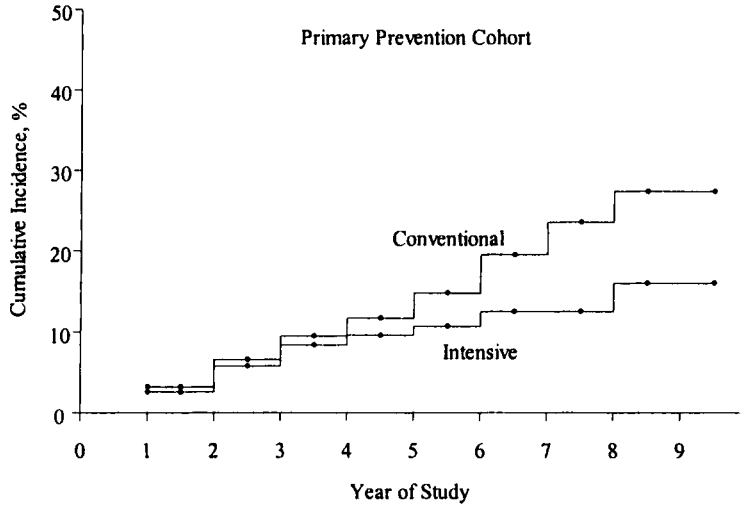
the lifetime exposure to hyperglycemia, or possibly related to some environmental exposure or characteristic.

Since the discovery of insulin in the 1920s, one of the principal issues of contention in the scientific community is what was often called the Glucose Hypothesis. This hypothesis asserts that the extent of exposure to elevated levels of blood glucose or hyperglycemia is the dominant determinant of the risk of diabetic nephropathy and other microvascular abnormalities or complications of type 1 diabetes. Among the first studies to suggest an association was a large observational study conducted by Pirart (1978a, 1978b) in Belgium over the period 1947–1973. This study examined the association between the level of blood glucose and the prevalence (presence or absence) of nephropathy. The data were obtained from a retrospective examination of the clinical history of 4,400 patients treated in a community hospital over a period of up to 25 years in some patients. The rather crude analysis consisted of figures that displayed the prevalence of nephropathy by year of diabetes duration for subgroups categorized as being in good, fair or poor control of blood glucose levels. These figures suggest that as the mean level of hyperglycemia increases, the risk (prevalence) of nephropathy also increases. This type of study is clearly open to various types of sampling or selection biases. Nevertheless, this study provides evidence that hyperglycemia may be a strong risk factor, or is associated with the risk of diabetic nephropathy. Note that this study is not strictly a prospective cohort study because the cohort was identified later in time and the longitudinal observations were then obtained retrospectively.

In all of these studies, biochemical measures of renal function are used to assess the presence and extent of nephropathy. Ultimately, however, end stage renal disease is characterized by the physiologic destruction of the kidney, specifically the glomeruli, which are the cellular structures that actually perform the filtration of blood. However, the only way to determine the physical extent of glomerular damage is to conduct a morphologic evaluation of a tissue specimen obtained by a needle biopsy of the kidney. As an example of a case-control study, Chavers, Bilous, Ellis, et al. (1989) conducted a retrospective study to determine the association between established nephropathy or not (the cases vs. controls) and evidence of morphologic (structural tissue) abnormalities in the kidneys (the risk factor or exposure). They showed that approximately 69% of patients with nephropathy showed morphologic abnormalities versus 42% among those without nephropathy, for a relative risk (odds ratio) of 3.2. Other studies (*cf.* Steffes, Chavers, Bilous and Mauer (1989) show that the earliest stage of nephropathy, microalbuminuria (which they defined as an AER $\geq$ 20 mg/24 h) is highly predictive of progression to proteinuria, with a positive predictive value ranging from 83–100%. These findings established that proteinuria is indeed associated with glomerular destruction and that microalbuminuria is predictive of proteinuria. Thus a treatment that reduces the risk of microalbuminuria can be expected to reduce the risk of progression to proteinuria, and one that reduces the risk of proteinuria will also reduce the extent of physiologic damage to the kidneys.

The major question to be addressed, therefore, was whether the risk of albuminuria or nephropathy could be reduced by a treatment that consistently lowered

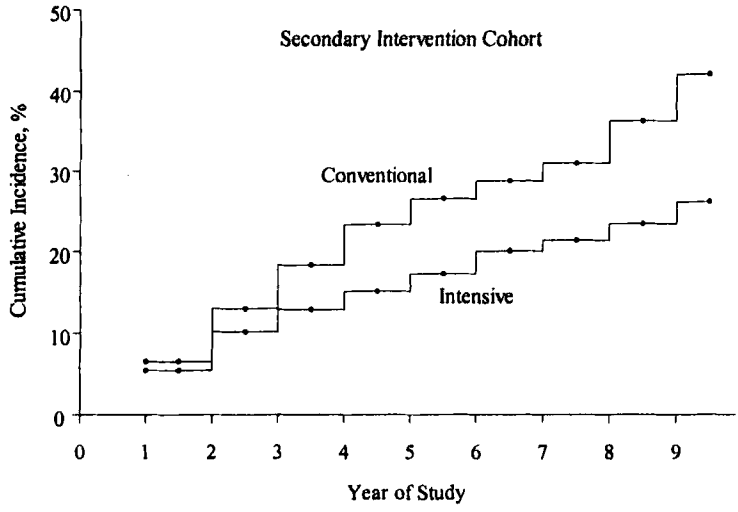
Fig. 1.1 Cumulative incidence of microalbuminuria (AER > 40 mg/24 h) over nine years of follow-up in the DCCT Primary Prevention Cohort.



the levels of blood glucose. By the 1980s, technological developments made an experiment (clinical trial) to test this hypothesis feasible. The level of blood glucose varies continuously over the 24 hour period, with peaks following meals and troughs before meals. It was discovered that the hemoglobin (red cells) in the blood become glycosylated when exposed to blood glucose. Thus the percent of the total hemoglobin that has become glycosylated (the HbA_{1c} %) provides an indirect measure of the mean level of hyperglycemia over the preceding 4–6 weeks, the half-life of the red blood cell. This made it possible to assess the average extent of hyperglycemia in individual patients. Other developments then made it possible for patients and their health-care teams to control their blood sugar levels so as to lower the level of hyperglycemia, as reflected by the level of HbA_{1c}. Devices for self-blood glucose monitoring allowed patients to measure the current level of blood glucose (mg/dL) from a drop of blood obtained by a finger prick. Patients could then alter the amount of insulin administered to keep the level of blood glucose within a desirable range. Also, a variety of types of insulin were developed, some of which acted quickly and some over long periods of time, that could be administered using multiple daily insulin injections or a pump. The health care team could then try different algorithms to vary the amount of insulin administered in response to the current level of blood glucose.

With these advances, in 1981 the National Institute of Diabetes, Digestive and Kidney Disease launched the Diabetes Control and Complications Trial (DCCT) to

Fig. 1.2 Cumulative incidence of microalbuminuria (AER > 40 mg/24 h) over nine years of follow-up in the DCCT Secondary Intervention Cohort.



test the glucose hypothesis (DCCT 1990, 1993). This was a large scale randomized controlled clinical trial involving 1441 patients enrolled in 29 clinical centers in the United States and Canada and followed for an average of 6.5 years (4–9 years). Of these, 726 patients comprising the primary prevention cohort were free of any microvascular complications (AER $\leq$ 40 mg/dL and no retinopathy, among other features); and 715 patients comprising the Secondary Intervention Cohort may have had minimal pre-existing levels of albuminuria (AER < 200 mg/dL) and mild retinopathy. Patients were randomly assigned to receive either intensive or conventional treatment. Intensive treatment used all available means (self-monitoring four or more times a day with three or more multiple daily injections or a pump in conjunction with diet and exercise) to obtain levels of HbA_{1c} as close as possible to the normal range (< 6.05%) while attempting to avoid hypoglycemia. Hypoglycemia occurs when the blood glucose level is reduced below a physiologically safe level, resulting is dizziness and possibly coma (unconsciousness) or seizures. Conventional treatment, on the other hand, consisted of one or two daily injections of insulin with less frequent self-monitoring with the goal of maintaining the clinical well-being of the patient, but without any specific glucose targets.

Figure 1.1 presents the cumulative incidence of microalbuminuria (AER > 40 mg/24 h) among the 724 patients free of microalbuminuria at baseline in the primary cohort (adapted from DCCT, 1993); presented with permission). The average hazard ratio for intensive versus conventional treatment (I:C) over the 9 years is 0.66. This