

Biomarkers: Potential Uses and Limitations

Richard Mayeux

Gertrude H. Sergievsky Center and the Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Summary: Biomarkers provide a dynamic and powerful approach to understanding the spectrum of neurological disease with applications in observational and analytic epidemiology, randomized clinical trials, screening and diagnosis and prognosis. Defined as alterations in the constituents of tissues or body fluids, these markers offer the means for homogeneous classification of a disease and risk factors, and they can extend our base information about the underlying pathogenesis of disease. Biomarkers can also reflect the entire spectrum of disease

from the earliest manifestations to the terminal stages. This brief review describes the major uses of biomarkers in clinical investigation. Careful assessment of the validity of biomarkers is required with respect to the stage of disease. Causes of variability in the measurement of biomarkers range from the individual to the laboratory. Issues that affect the analysis of biomarkers are discussed along with recommendations on how to deal with bias and confounding. **Key Words:** Antecedent biomarkers, diagnostic biomarkers, variability, reliability, validity.

INTRODUCTION

Biological markers (biomarkers) have been defined by Hulka and colleagues¹ as “cellular, biochemical or molecular alterations that are measurable in biological media such as human tissues, cells, or fluids.” More recently, the definition has been broadened to include biological characteristics that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.² In practice, biomarkers include tools and technologies that can aid in understanding the prediction, cause, diagnosis, progression, regression, or outcome of treatment of disease. For the nervous system there is a wide range of techniques used to gain information about the brain in both the healthy and diseased state. These may involve measurements directly on biological media (e.g., blood or cerebrospinal fluid) or measurements such as brain imaging which do not involve direct sampling of biological media but measure changes in the composition or function of the nervous system.

Biomarkers of all types have been used by generations of epidemiologists, physicians, and scientists to study human disease. The application of biomarkers in the diagnosis and management of cardiovascular disease, in-

fections, immunological and genetic disorders, and cancer are well known.^{1,3} Their use in research has grown out of the need to have a more direct measurement of exposures in the causal pathway of disease that is free from recall bias, and that can also have the potential of providing information on the absorption and metabolism of the exposures.⁴ Neuroscientists have also relied on biomarkers to assist in the diagnosis and treatment of nervous system disorders and to investigate their cause. Blood, brain, cerebrospinal fluid, muscle, nerve, skin, and urine have been employed to gain information about the nervous system in both the healthy and diseased state. This paper focuses on biomarkers as defined by Hulka et al.,¹ i.e., direct measures of biological media, and other papers in this issue will address brain imaging and other markers.

The rapid growth of molecular biology and laboratory technology has expanded to the point at which the application of technically advanced biomarkers will soon become even more feasible.^{5–8} Molecular biomarkers will, in the hands of clinical investigators, provide a dynamic and powerful approach to understanding the spectrum of neurological disease with obvious applications in analytic epidemiology, clinical trials and disease prevention, diagnosis, and disease management.

TYPES OF BIOMARKERS

Biomarkers have been classified by Perera and Weinstein³ based on the sequence of events from exposure to

Address correspondence and reprint requests to Dr. Richard Mayeux, Sergievsky Center, 630 West 168th Street, New York, New York 10032. E-mail: rpm2@columbia.edu.

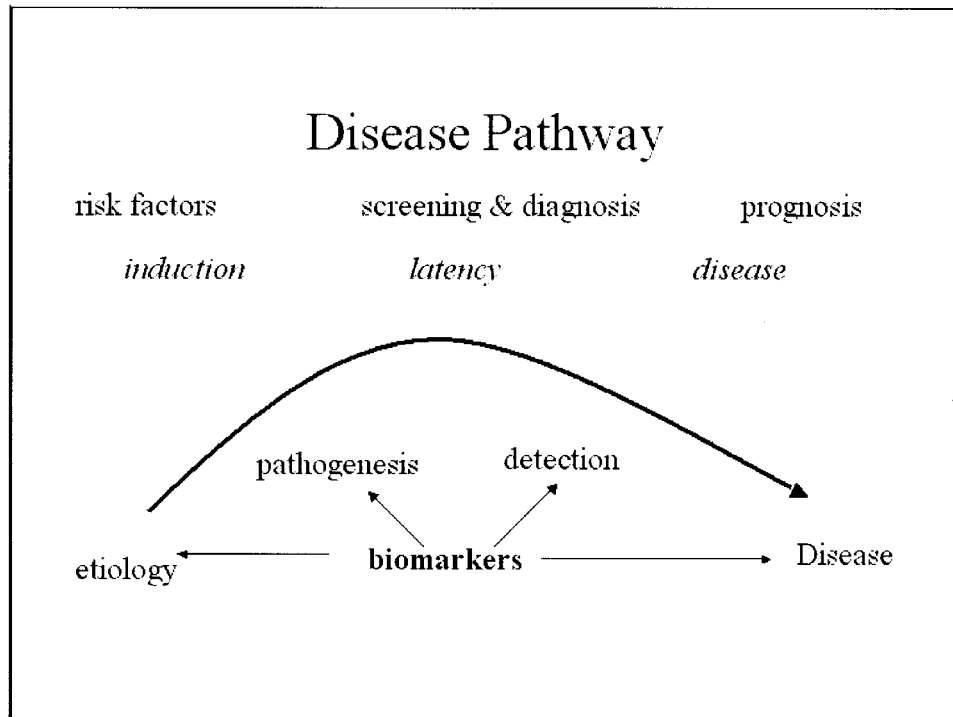


FIG. 1. Disease pathway and potential impact of biomarkers.

disease (FIG. 1). Though biomarkers readily lend themselves to epidemiological investigations, they are also useful in the investigation of the natural history and prognosis of a disease. Schulte⁹ has outlined the capabilities of biomarkers (Table 1). In addition to delineating the events between exposure and disease, biomarkers have the potential to identify the earliest events in the natural history, reducing the degree of misclassification of both disease and exposure, opening a window to potential mechanisms related to the disease pathogenesis, accounting for some of the variability and effect modification of risk prediction. Biomarkers can also provide insight into disease progression, prognosis, and response to therapy.

There are two major types of biomarkers: biomarkers

TABLE 1. Contributions of Valid Biomarkers to Clinical Research

Capabilities of Biomarkers

- Delineation of events between exposure and disease
- Establishment of dose-response
- Identification of early events in the natural history
- Identification of mechanisms by which exposure and disease are related
- Reduction in misclassification of exposures or risk factors and disease
- Establishment of variability and effect modification
- Enhanced individual and group risk assessments

Modified from Schulte.⁹

of exposure, which are used in risk prediction, and biomarkers of disease, which are used in screening and diagnosis and monitoring of disease progression. Biomarkers used in risk prediction, in screening, and as diagnostic tests are well established, and they offer distinct and obvious advantages. The classification of many neurological diseases is based on either standardized clinical criteria or histological diagnoses. Biomarkers also have the potential to identify neurological disease at an early stage, to provide a method for homogeneous classification of a disease, and to extend our knowledge-base concerning the underlying disease pathogenesis. These advantages have direct application to all types of clinical investigation, from clinical trials to observational studies in epidemiology.

In epidemiological (or quasi-experimental) investigations, biomarkers improve validity while reducing bias in the measurement of exposures (or risk factors) for neurological disease. Rather than relying on a history of exposure to a putative risk factor, direct measurement of the level of exposure or the chromosomal alteration resulting from the exposure lessens the possibility of misclassification of exposure. Such misclassifications not only produce inaccurate and deceptive results but also reduce the power of studies to detect health effects. Thus, the use of biomarkers improves the sensitivity and specificity of the measurement of the exposures or risk factors.

Molecular biomarkers have the additional potential to identify individuals susceptible to disease.⁶ Molecular

genetics have already had an impact on neurological practice, leading to improved diagnosis. Classification of populations in terms of the degree of susceptibility on the basis of such biomarkers produces greater accuracy than relying on historical definitions of susceptibility.^{7,10,11} For example, a biomarker will allow the stratification of a population on the basis of a specific “genotype” associated with a disease rather than relying on a report of the “family history” of the disease. The ability to quantify “susceptibility” in this way can be an extremely important method for estimating disease risk among various populations.

BIOMARKERS OF EXPOSURE OR ANTECEDENT BIOMARKERS

Environmental exposures, effect modifiers, or risk factors

When a disease is suspected of resulting from a toxic exposure, researchers naturally wish to measure the degree of exposure. External exposure is the measured concentration of the toxin in an individual’s immediate environment. While questionnaires offer an historical account of the exposure, direct measurement of the alleged toxin in the air, water, soil, or food can provide accurate information regarding the “dose” of the exposure. Measurement of the external dose provides the basis to understand the relationship to the disease process, but a measurement of “internal” dose may provide more accuracy.

When the toxin is identified in tissues or body fluids it becomes a biomarker for the internal dose. A biomarker that measures a “biologically effective dose” generally indicates the amount of toxin or chemical measured in the target organ or its surrogate. Lead exposure is an excellent example. A history of lead exposure can be strengthened by measurement of lead in the environment, but the best indication of the dose of exposure may be determined in blood and tissues (hair, nails, teeth). The pharmacokinetic properties of the toxin or chemical of interest becomes important to consider in measurement of the internal dose because a number of body fluids could be used based on the pharmacologic properties of the agent. Some chemicals such as halogenated hydrocarbons are stored in adipose tissue but others, such as organophosphate pesticides, are better measured in blood or urine.

Most biomarkers of exposure measure antecedent factors thought to modify (increase or decrease) the risk of developing the disease investigated, as illustrated in Figure 1. The advantage of a biomarker of exposure over a history of exposure is that it estimates the actual “internal” dose of the exposure. This improves precision in the measurement of any risk factor by adding both internal and external validity when examining the effect of the exposure on the outcome. Biomarkers are particularly

useful in the cross-sectional investigation of acute disease because of the pharmacologic properties of the chemical or toxin. It is very difficult to find biomarkers for exposures that are stable over the long periods required for prospective studies of chronic neurological diseases such as Alzheimer’s disease. Banked serum or plasma may be of value in some instances depending on the disorder being investigated and the pharmacologic characteristics of the biomarker. Issues of timing, persistence, dose, and storage site all must be considered for this class of biomarker.

Genetic susceptibility

Epidemiologic analyses can examine familial aggregation and assess genetic and environmental contributions to a disease by using life table methods and recurrence risk. Mutations in genes that result in Mendelian forms of disease are typically deterministic. Variant alleles in genes or polymorphisms may be related to susceptibility but are not deterministic. Most adult-onset degenerative diseases of the nervous system are likely to be a composite of related characteristics, heritable and environmental. The correlated combinations of these features constitute the trait or disease. Therefore, these types of antecedent biomarkers may or may not be directly involved in the etiology. In some instances the genetic variant is neither necessary nor sufficient to cause the disease. However, they can be powerful antecedents at any stage of the disease pathway, as illustrated in Figure 1. By definition these antecedent biomarkers exist before the disease or the outcome occurs and are independent of other exposures. They improve the precision in the measurement of other associations, because they may be synergistic or antagonistic.

For neurological disorders, biomarkers of genetic susceptibility are rapidly becoming more available. Identification of the variant allele in a gene, such as *APOE* (apolipoprotein E), is quite useful in assessing risk and in providing information regarding the pathogenesis of the Alzheimer’s disease. With this information investigators can now examine other genes or environmental risk factors to determine whether they modified (increase or decrease) the risk of Alzheimer’s disease. Similarly, variations in several genes appear to influence susceptibility to Parkinson’s disease, which has also been related to environmental risk factors. Once established, a specific genotype might be used to predict an association with a particular environmental toxin.¹²

Intermediate biomarkers

Some biomarkers represent direct steps in the causal pathway of a disease and are therefore strongly related to disease. Others are related in some indirect way to the cause. There are numerous possibilities to consider. A biomarker could be dependent on another known or unknown factor to cause disease. Thus, it is not the only

determinant but it is in the causal pathway and remains strongly related to the disease. The biomarker could also be related to an exposure that has already been identified or represents an alteration caused by the exposure that results in the disease. The most precarious situation is one in which the biomarker is related to some unknown factor that is also related to the exposure. This type of confounder, if unidentified, can decrease the validity of the association between the biomarker and the disease.

BIOMARKERS OF DISEASE

Screening, diagnostic tests, and prognosis

Biomarkers depicting prodromal signs enable earlier diagnosis or allow for the outcome of interest to be determined at a more primitive stage of disease. Blood, urine, and cerebrospinal fluid provide the necessary biological information for the diagnosis. In these conditions, biomarkers are used as an indicator of a biological factor that represents either a subclinical manifestation, stage of the disorder, or a surrogate manifestation of the disease. Biomarkers used for screening or diagnosis also often represent surrogate manifestations of the disease. The potential uses of this class of biomarkers include: 1) identification of individuals destined to become affected or who are in the "preclinical" stages of the illness, 2) reduction in disease heterogeneity in clinical trials or epidemiologic studies, 3) reflection of the natural history of disease encompassing the phases of induction, latency and detection, and 4) target for a clinical trial. The improvement in validity and precision far outweigh the difficulty in obtaining such tissues from patients.

Most ethical review boards and the healthcare systems require adequate follow-up for individuals that screen positive regardless of whether or not they have the disease. Also, treatment should be available for those who screen positive and it must be accessible and acceptable. Those who screen positive and are diseased should be allowed access to treatments and those treatments must be adequate and available. It is useful to remember that the main benefit of screening is primary (before onset of symptoms) or secondary (early or prodromal detection) prevention. Consider the benefits of conducting a therapeutic trial in patients before overt manifestations occur.

Diagnostic tests for neurological diseases are used with increased frequency in clinical research and practice. In the diagnostic effort, collection of information from various sources, some of which includes results from diagnostic tests, helps to achieve the ultimate goal of increasing the probability of a given diagnosis. Clinical tests are also performed, though probably less often, for other reasons such as the following: to measure disease severity, to predict disease occurrence, or to monitor the response to a particular treatment. More importantly, biomarkers for disease easily lend themselves to clinical

trials. Another advantage of this type of diagnostic test is the reduction in disease heterogeneity in clinical trials or observational epidemiologic studies, leading to better understanding of natural history of disease encompassing the phases of induction, latency and detection.

VARIABILITY

Although biomarkers have numerous advantages, variability is a major concern. Variability applies regardless of whether the biomarker represents an exposure or effect modifier, a surrogate of the disease, or an indication of susceptibility. Interindividual variability can result from the amount of an external exposure or from the way a putative toxin is metabolized. For example, individuals exposed to the same chemical might differ in their ability (or inability) to metabolize the agent, or they may have experienced different types of exposures (in the field as compared with in the office). Intraindividual variability is usually related to laboratory errors or other conditions, or exposures unique to the individual. Group variability is also encountered, but this is often the desired outcome of a study. Obviously, it is best when group differences are large. Nonetheless, the ability of a biomarker to distinguish between groups is measured by sensitivity and specificity or similar variance estimates. Consideration of the sources of variability in the measurement of a biomarker decreases the potential for misclassification of the exposure.

While measurement error is always a concern with biomarkers, other important factors may explain individual or group variability. Some workers may always wear protective equipment whereas others may not. Interaction with other exposures, drugs, or effect modifiers can increase or decrease the effect of the biomarker under consideration as an exposure or as a measure of susceptibility. Variability can also be attributed to the effects of factors such as individual diet or other personal characteristics. The amount of dietary fat can influence the biological measurement of lipid-soluble vitamins as well as toxic chemicals. These individual factors must be considered by the investigator to fully establish the major causes of variability in these investigations.

VALIDITY

Precise numbers are enticing, but they are prone to the same problems as any variable. Reliability, validity, sensitivity, specificity, ascertainment bias, and interpretation of data using biomarkers should be reviewed just as carefully as any other variable. These problems remain whether the biomarker is being used as a variable in a clinical trial or in an epidemiologic study.

Reliability or repeatability is crucial. Laboratory errors can lead to misclassification of exposures or disease if

the biomarker is not reliable. Pilot studies should be performed to establish a reasonable degree of reliability. Changes in laboratory personnel, laboratory methods, storage, and transport procedures may all affect the reliability of the biomarkers used in any investigation. Kappa statistics for binary or dichotomous data and intraclass correlation coefficients should be used to assess test–retest agreement and consistency.

The evaluation of the validity of a biomarker is complex. Schulte and Perera¹³ suggest three aspects of measurement validity: 1) content validity, which shows the degree to which a biomarker reflects the biological phenomenon studied, 2) construct validity, which pertains to other relevant characteristics of the disease or trait, for example other biomarkers or disease manifestations, and 3) criterion validity, which shows the extent to which the biomarker correlates with the specific disease and is usually measured by sensitivity, specificity, and predictive power.⁴ To further evaluate the effect of misclassification of disease, false positives and false negatives as well as positive and negative predictive power should also be estimated. In an ideal situation the biomarker has a clear predictive value but in many cases one needs to be established. The use of receiver–operator characteristic curves can provide the tools necessary to determine the best choice in terms of sensitivity and false-positive rates, particularly when other tests are used.^{14,15}

Most would agree that screening tests would be very desirable for chronic progressive disorders. One purpose of screening is early detection with the hope of preventing the illness altogether. Many of the methods and concerns related to diagnostic testing apply to screening as well. As with other diagnostic methods, sensitivity and specificity tell us the accuracy of the test but not the probability of disease. For that we need to estimate the predictive values (positive and negative). Positive predictive value (PPV) is the percentage of people with a positive test who actually have the disease. This provides us with information about the likelihood of the disease being present if the test is positive. Negative predictive value (NPV) is the percentage of people with a negative test who do not have the disease. Increasing the prior probability will increase the PPV but decrease the NPV, assuming that the sensitivity and specificity remain unchanged. Similar changes in the predictive values occur with changes in the prevalence of a condition as will be discussed in screening.

Since validity is measured by sensitivity and specificity and predictive power by PPV and NPV, a major difference in evaluating screening and diagnostic tests is the pretest probability. Screening, by definition, includes a larger number of individuals without the disease, generally ascertained via a defined population sample. Diagnostic tests are designed to improve clinical diagnoses by enhancing the probability of disease, and by definition

TABLE 2. *Advantages and Disadvantages of Biomarkers*

Advantages	Disadvantages
Objective assessment	Timing is critical
Precision of measurement	Expensive (costs for analyses)
Reliable; validity can be established	Storage (longevity of samples)
Less biased than questionnaires	Laboratory errors
Disease mechanisms often studied	Normal range difficult to establish
Homogeneity of risk or disease	Ethical responsibility

the pretest probability would be high. However, for screening the prior probability is much lower and that effect will lower the PPV. Therefore, screening also requires careful consideration of prevalence, or the prior probability of disease. These analytic methods are now available on many software statistical packages.

The investigator must be clear about the use of the biomarker in the study. Errors are most often made when biomarker data are over interpreted. For example, the results of one study may indicate that a specific biomarker (collected as a measure of an exposure or susceptibility) is strongly associated with a particular disease or outcome. The investigator, on the other hand, interprets the result as a biomarker for the disease or the observed outcome. No matter how high the odds ratio or relative risk, a biomarker of this type could not be expected to function as a diagnostic test unless it is a manifestation of the disease. For example, the APOE- $\epsilon 4$ allele is strongly associated with Alzheimer's disease, but its presence does not infer disease. Many patients without an APOE- $\epsilon 4$ allele develop Alzheimer's disease and some individuals with an APOE- $\epsilon 4$ allele do not develop this condition.¹⁶

PRACTICAL CONSIDERATIONS AND CONCERNS

The advantages and disadvantages of biomarkers are shown in Table 2. Several points should be considered before using biomarkers in clinical studies of any type.

Measurement errors

Imperfect measurement of the biomarker would naturally lead to decreased validity of the relation to the disease. However, there are numerous types of measurement errors other than those errors that occur in the laboratory. Problems with the collection equipment or in the transportation of specimens to the laboratory can affect the measurement of the biomarker. Improper storage of samples or changes in storage environment can also affect measurement of biomarkers. Technicians are

the handlers of most specimens and so appropriate training of new personnel is essential. Finally, receipt and control errors such as in the transcription of identification numbers if done by hand can always be source of error. A well organized procedures manual outlining the details for documentation, storage, monitoring of specimens and maintaining records, can alleviate many of these issues. Most laboratories and large-scale studies institute a quality assurance and quality-control program to reduce measurement errors.

Bias

Bias occurs in any study including those with biomarkers. When biases occur without regard to the outcome, so-called nondifferential bias, the effects on the study are less serious but favor the null hypothesis of no association. Problems arise when availability of the biomarker is differentially related to either the disease or the exposure or when the specimen acquisition, storage, measurement, or ascertainment procedures differ in those with the disease compared to those without the disease or outcome of interest. Differential biases tend to favor an association in either direction, which may not be the true relationship between the biomarker and the disease. To reduce such biases, a high response rate from all cases and controls should be maintained and the investigators should have an objective review board review and monitor the conduct of the study, observing possible biases in subject participation or specimen ascertainment.

Confounding

The most important source of confounding is the failure to identify factors that may alter the measurement of the biomarker. These can be internal, such as the weight of the subject, or external, such as the batch of laboratory kits used. Individual properties of biomarkers should influence the choice and interpretation for its inclusion in any investigation. The effects of potential confounders such as age, gender, diet, and other metabolic factors should be investigated before initiating the investigation. Biologic stability is critical particularly if the biomarker is to be stored for any length of time. Banked serum or plasma is of great value in any study unless it affects the pharmacologic properties of the biomarker. For example, some nutrients such as vitamins do not store well because they are light-sensitive. Storage of all tissues including lymphocytes and extracted DNA can be expensive and the stability of the biomarker studies must be evaluated if storage is required for long intervals. These are often overlooked in the analyses and can seriously affect the outcome. One should use data on potential confounders when designing the study and collect relevant internal and external information that might effect the measurement. This information can be included in the analysis of the relation between the biomarker and the outcome of interest.

Cost

The choice of the biomarker for research should be guided by the scientific question and by the financial resources. Cost is always a concern. In a small clinical trial this may be important; if an epidemiologic study includes thousands of subjects the cost can be quite high unless the laboratory procedure is automated and relatively simple. In fact, for some investigations larger sample sizes can bring down the cost per subject. This generally implies that the biomarker is readily available and its inclusion in the study is feasible. For example, automated procedures have made the inclusion of lipid profiles in clinical studies of stroke quite feasible. Methods have improved to the point that a "finger-stick" can provide the necessary amount of blood. Depending on the type of investigation, researchers should have an idea of the false-positive or false-negative profile of the biomarker. As might be expected "false positives" create extra work regardless of whether it is a biomarker of exposure, susceptibility, or disease. "False negatives" simply increase the overall cost of the study. Tolerance for this problem depends on the funding available.

Acceptability

Because biomarkers are derived from human tissues or body fluids, the choice of biomarkers is not trivial. Biomarkers can be also associated with some degree of risk. In clinical trials, this is less a concern because the patient will possibly benefit from the "new treatment." In quasi-experimental studies, the source of the biomarker may be critical. Body fluids such as blood and urine are usually well tolerated. However, biopsy (particularly of neural tissue) and collection of cerebrospinal fluid are more difficult and associated with slight risks. Risk-benefit will be an issue for the investigator to resolve. Pilot studies are always quite helpful for convincing institutional review boards that your study is safe and that the risk-benefit ratio favors a benefit.

CONCLUSIONS

Many studies using biomarkers never achieve their full potential because of the failure to adhere to the same rules that would apply for the use of variables that are not biological. The development of any biomarker should precede or go in parallel with the standard design of any epidemiological project or clinical trial. In forming the laboratory component, pilot studies must be completed to determine accuracy, reliability, interpretability, and feasibility. The investigator must establish "normal" distributions by important variables such as age and gender. The investigator will also want to establish the extent of intraindividual variation, tissue localization, and persistence of the biomarker. Moreover, he or she will need to determine the extent of interindividual variation attribut-

able to acquired or genetic susceptibility. Most, if not all of these issues can be resolved in pilot studies preceding the formal investigation.

Acknowledgments: This work was supported by National Institutes of Health Grants AG07232, AG08702, the Charles S. Robertson Memorial Gift for Alzheimer's Disease Research from the Banbury Fund, and the Blanchette Hooker Rockefeller Foundation.

REFERENCES

- Hulka BS. Overview of biological markers. In: Biological markers in epidemiology (Hulka BS, Griffith JD, Wilcosky TC, eds), pp 3–15. New York: Oxford University Press, 1990.
- Naylor S. Biomarkers: current perspectives and future prospects. *Expert Rev Mol Diagn* 3:525–529, 2003.
- Perera FP, Weinstein IB. Molecular epidemiology: recent advances and future directions. *Carcinogenesis* 21:517–524, 2000.
- Gordis L. Epidemiology and public policy. In: Epidemiology (Gordis L, ed), pp 247–256. Philadelphia: W.B. Saunders, 1996.
- Verbeek MM, De Jong D, Kremer HP. Brain-specific proteins in cerebrospinal fluid for the diagnosis of neurodegenerative diseases. *Ann Clin Biochem* 40:25–40, 2003.
- Galasko D. New approaches to diagnose and treat Alzheimer's disease: a glimpse of the future. *Clin Geriatr Med* 17:393–410, 2001.
- Rohlf C. Proteomics in neuropsychiatric disorders. *Int J Neuropsychopharmacol* 4:93–102, 2001.
- Reiber H, Peter JB. Cerebrospinal fluid analysis: disease-related data patterns and evaluation programs. *J Neurol Sci* 184:101–122, 2001.
- Schulte PA. A conceptual and historical framework for molecular epidemiology. In: Molecular epidemiology: principles and practices (Schulte PA, Perera FP, eds), pp 3–44. San Diego: Academic Press, 1993.
- Merikangas K. Genetic epidemiology: bringing genetics to the population—the NAPE Lecture 2001. *Acta Psychiatr Scand* 105:3–13, 2002.
- Muller U, Graeber MB. Neurogenetic diseases: molecular diagnosis and therapeutic approaches. *J Mol Med* 74:71–84, 1996.
- Mayeux R. Epidemiology of neurodegeneration. *Annu Rev Neurosci* 26:81–104, 2003.
- Schulte PA, Perera FP. Validation. In: Molecular epidemiology: principles and practices (Schulte PA, Perera FP, eds), pp 79–107. San Diego: Academic Press, 1993.
- Pepe MS, Thompson ML. Combining diagnostic test results to increase accuracy. *Biostatistics* 1:123–140, 2000.
- Thompson ML, Zucchini W. On the statistical analysis of ROC curves. *Stat Med* 8:1277–1290, 1989.
- Mayeux R, Saunders AM, Shea S et al. Utility of the apolipoprotein E genotype in the diagnosis of Alzheimer's disease. Alzheimer's Disease Centers Consortium on apolipoprotein E and Alzheimer's disease. *N Engl J Med* 338:506–511, 1998.