APPROPRIATE USE OF POPULATION AND INDIVIDUAL REFERENCE DATA IN PATIENT CARE

Eugene K. Harris, Ph.D.

Division of Computer Research and Technology, National Institutes of Health, Bethesda, Maryland 20892

Abstract

This report gives a brief description of two general areas of statistical research with application to the management of clinical laboratory data: 1) Time series models for short series of observed values from a single patient; 2) Empirical Bayes methods to improve estimates of individual patient means, using population reference data.

It is a cliche in the modern clinical laboratory that laboratory directors and their clinical colleagues are being inundated with data concerning the biochemical and hematological status of patients. While outdated, non-specific laboratory tests are being dropped at an arithmetic rate, new biochemistries are being added at an exponential rate. The credit or blame for this situation can be awarded primarily to the rapid pace of research in clinical biochemistry coupled with equally rapid improvements in the chemist's ability to measure small concentrations of important blood constituents, such as enzymes, hormones, vitamins, and free ions. In addition, the pressures of manufacturers to market high-speed analyzers; of lawyers, to prosecute medical malpractice suits; and of doctors themselves to keep up with their peers, all force increasing volumes of numerical data to flow from hospital laboratories. Unfortunately, improvements in our knowledge of physiological processes and the evolution of disease come much more slowly. Moreover, while many non-clinical scholars (biochemists, computer scientists, electronics engineers, even statisticians) contribute to the creation and management of clinical laboratory data, only the physician is licensed to apply this mass of numerical information to its ultimate purpose: care of the sick. The result is predictable: indigestion. Many physicians resolve this problem by applying a special kind of preventive medicine -- they simply ignore the data set before them, or else explain it all away as confirming a diagnosis already determined by traditional means.

Now, for decades physicians have had available to them, a statistical "crutch" to support this sweepout of laboratory data -- namely, the "normal range", a mysterious interval attached to each constituent, such that, if the measured value of that constituent fell within this range, it could safely be ignored, while if it fell outside, it probably was either a laboratory error or was expected from the physician's pre-existing knowledge of the case. Normal ranges are still with us, and in ever greater abundance. Instead of deriving from measurement of ten or twenty "good old days", they now arise from observation of many hundreds of presumably healthy individuals, sometimes even drawn by probability sampling and almost always classified by age and sex. However, the use of such normal ranges in routine medical practice, have not changed.

Within the last decade, collaboration between statisticians and clinical pathologists has begun to shake, just slightly, the pre-eminence of the population-based normal range in governing the doctor's interpretation of laboratory data. Many recent studies have demonstrated that what most physicians long recognized intuitively: that a healthy individual, under reasonably controlled conditions of sampling, usually exhibits much less variability over time in his blood chemistries and hematoelges than the normal range indicates. In other words, between-person variation, even among healthy persons, greatly exceeds within-person variation. This is particularly true for the enzymes and hormones with whose measurement the modern clinical laboratory is increasingly concerned.

The significance of these findings, particularly for programs of preventive medicine, is obvious. The population-based normal range of most laboratory tests is insensitive to real deviations from an individual's past record of observed values. However, to isolate real departures from random, within-person variation requires developing simple statistical models appropriate for short, stationary and nonstationary time series. Most of my own work during the past few years has been devoted to the study of models of within-person variability and the development of computer programs to implement them in clinical laboratories.*

These methods require at least three periodic measurements of an individual patient before they can be applied. But before there are three observations, there will be only one, and then two, and medical decisions will be required at these points. Are normal ranges still the only recourse here? If the answer were yes, then we would be ignoring the individuality represented in even a single measurement from a specific patient. Again, statistical theory has shown the way. We make use of population data, but only to improve the estimate of the individual's true mean which has been provided by the one or two

U.S. Government work not protected by U.S. copyright.
measurements obtained. Note that the essential purpose of the time series model applied to three or more past measurements was to improve sensitivity, i.e. to increase the physician's ability to recognize a nonrandom change. In contrast, the use of population data in connection with a single measurement will be to improve specificity, i.e. to reduce the well-recognized danger of false positive diagnosis based on a single value which happens to lie outside the "normal range".

The statistical methods referred to in this situation come from Empirical Bayesian estimation theory.4-7 The population data represents the "prior" probability distribution; hence, the adjective, Empirical. The new estimate is the mean of the "posterior" distribution of the patient's true value, computed as a weighted average of the patient's measurement and the mean of the reference population. The weighting coefficient is the ratio of expected within-person to between-person variation. Thus, if individuality in a particular blood constituent is known to be high, little change will be made in the observed value. Apart from this, there are further safeguards which may be employed (e.g.: "limited translation estimators") to protect an apparently aberrant observation from too close conformity to the reference distribution. The end effect, however, is to reduce the probability of a false positive report based on a single, highly deviant observation. Of course, the average of two independent observations would be subjected to a smaller adjustment since it represents a more precise estimate of the patient's true value.

All these methods, time series models for detecting nonrandom change and Empirical Bayesian formulas for improving the estimation of true means, may be extended to multivariate databases, e.g. batteries of laboratory tests for studying a single organ, or the various hematologic measurements obtained from a Coulter S analyzer. More importantly, these programs should really be incorporated in an interactive laboratory data-processing system capable of producing cumulative reports immediately on request for any patient receiving care. Of course, much more experience is needed to ascertain the behavior of these methods in routine use in hospital or outpatient clinic. In support of "management by exception", they offer a statistical screen, more valid and, I hope, more informative than the "normal range", which the clinical laboratory director and his ward colleagues may find useful in coping with the ever-increasing flow of laboratory data.

EUGENE K. HARRIS, PH.D.
DIVISION OF COMPUTER RESEARCH AND TECHNOLOGY
NATIONAL INSTITUTES OF HEALTH
BUILDING 12A, ROOM 2041
9000 ROCKVILLE PIKE
BETHESDA, MARYLAND 20014

References