Applications of the PROPHET system in human clinical investigation

by BERNARD J. RANSIL

Boston City Hospital
Boston, Massachusetts

INTRODUCTION

The study of chemical-biological interrelationships reaches its most challenging levels of complexity and difficulty in human clinical investigation. Whether the objective is development of a new drug, better understanding of an old one, or further elucidation of life processes which normally or pathologically distribute, transform, and eliminate exogenous and endogenous chemical substances, there is no experimental subject more important and more complex than man. And since no experimental subject is more difficult to study systematically than man, there probably is no area of biology and medicine where computer science and technology are needed more, but exploited less, than in human clinical investigation.

This paper surveys the nature of clinical investigative research and the information-handling problems associated with it. Selected topics are illustrated by describing how the PROPHET system† has been used in the Clinical Research Center of the Harvard Medical Unit at Boston City Hospital. Among other things, the discussion attempts to show that PROPHET possesses a sufficiently large and powerful repertoire of features to make it not only an invaluable tool in clinical pharmacology but also an instructive paradigm of a protocol-oriented, research data-handling system for almost any type of clinical investigation.

CLINICAL INVESTIGATION

Human clinical investigation is both descriptive and quantitative, inductive and deductive, concerned with both individual and collective events. Not unlike statistics (defined somewhere as “the art of being precise though vague”), it possesses an applicational elusiveness arising in part perhaps from the range and actuarial nature of the phenomena it explores.

Human clinical investigation begins with observations of various types and degrees of precision and accuracy at various organizational levels, concerning the function and malfunction of the human body, made upon identifiable subjects or groups of subjects for specified periods of time under specified circumstances. The less known about a phenomenon under study, usually the more descriptive the approach. As the phenomenon, process or mechanism under study approaches less complex levels of structural organization and function (organism—organ—tissue—molecule), the greater the potential for more quantitative studies, for more rigorous experimental control, and greater investigative precision.

The activity of clinical investigation depends on data—its collection, storage, retrieval and utilization as a basis for understanding disease processes in man, and their cure, reversal or control, in a context that can range from the so-called “controlled” to “non-controlled” investigative conditions.

The “controlled” experiment attempts to study a phenomenon in terms of both a perturbed and unperturbed state, or in terms of one state compared to some defined reference state. This may be done in a number of ways, which is a topic itself in experimental design.

The “controlled experiment” is demarcated in both time and population size. Its conclusions refer to the small samples studied but are often extended (with varying degrees of validity) to populations-at-large on the basis of sampling theory, under the assumption that a true cross-section obtains. Most modern clinical investigation is conducted in this mode.

Less common nowadays is that form of clinical investigation arising from the so-called “bedside observation.” As practiced by the early pioneers of clinical medicine, at a time when the practice of medicine was more personal, more descriptive, less precise, less academically oriented, this technique consisted of periodic observations of scores or hundreds of patients over time spans of months to years. It was...
inseparable from patient care and usually employed the long term qualitative and descriptive patient care data as its data base for comparison, correlation and inference. The "data bank" was one physician's files and personal experience; the "memory" his own; the primary data processing technique was "clinical correlation" using trial-and-error therapeutics with outcome (improvement, cure, no improvement, death, etc.) as a measure of therapeutic efficacy upon which the value of a drug, therapeutic regimen or accuracy of diagnosis was judged. Transmission of information was primarily by example, word of mouth and the published case history. This form of investigation—uncontrolled by modern standards, descriptive, qualitative for the most part, correlative—was open-ended as to time and population sample. Generalization to larger populations, always possessed of an element of risk even with the best controls and sampling techniques, possessed sufficient validity to produce recognizable improvements in patient diagnosis and treatment. In its own way, judging from the contribution it made to the understanding of many disease processes and the alleviation of human suffering, it worked.

CLINICAL INVESTIGATORS

Contemporary clinical investigation is performed by individuals, groups and teams possessing incredibly diverse backgrounds, training and levels of competence. They range from high school students on summer projects to multiple-degreed scientists with decades of experience in their specialties. They include individuals formally trained in medicine, the biological, physical, mathematical, engineering and social sciences.

Because most clinical investigation is rooted in the study of small representative samples of subjects or specimens, a working knowledge of statistics is mandatory. This includes a good knowledge of, and feel for, the notion of distributions (especially Gaussian, Poisson and binomial) and their central measures, the theory of sampling, error analysis and confidence limits, the basic idea of significance tests and the common types of significance tests for the most frequently encountered distributions, the notion of correlation for one or more variables, trend analysis, ranking tests, curve fitting and measures of goodness of fit, analysis of data for internal consistency and elimination of outliers.

It is probably not overstating the case to say that the acquisition of the prerequisite statistical, mathematical and computing skills is not as urgent a priority to most clinical investigators in their formal scientific training and formation as the development of expertise in the preclinical and clinical sciences. Having been exposed to the customary elementary course in statistics, usually taught by statisticians or mathematicians with cursory concern for applications, many become discouraged from further formal training by the calculational aspects and rely on picking up what they "need to know" for problem solving from colleagues and the research literature. Perusal of the clinical investigative literature reveals a tremendous variability of statistical rigor and facility not infrequently taxing credibility, despite an ever increasing demand from journal reviewers for better data reportage and analysis. It is the author's impression that many clinical investigators use biometrics not as an analytic tool to evaluate what the data actually may say about a phenomenon or a system, but as a set of required rules one follows grudgingly, mechanically, in follow-the-leader fashion to verify or rule out a working hypothesis.

DATA PROCESSING REQUIREMENTS

From the foregoing descriptions we may infer that the data processing capabilities required for clinical investigation include: collection, storage, retrieval, sorting, interfacing, transmission, analysis and diagnosis. Collection can be in several modalities—written anecdotal and descriptive data from human observers (both textual and numerical), digital or analogue data from instruments. Storage can be long or short term; if the former, the concept of a data bank arises. Data analysis can range over all forms of statistical analysis, and a variety of mathematical operations such as curve fitting and model building.

Given the nature of clinical investigation, its data processing and analysis requirements, and the general level of statistical-mathematical-computing competence of its practitioners, what characteristics might a data processing system ideally be expected to possess for clinical investigative applications?

Responses to this question can be as varied and unique as the investigators, applications and clinical research environments themselves. The author's experience in a specific clinical investigative setting has led him to identify a number of factors that must be considered. How one orders these factors as priorities to match his overall resources, constraints and goals will largely determine the system design and its hardware/software configurations. These factors include:

- Economy of initial cost, operation, maintenance and repairs, user time.
- Features important to non-specialist users: reliability—low down time, accessibility, flexibility, versatility, simplicity of operation, minimal or no programming, programmed instruction capability, interactive dialogue and batch processing modes.
- Hardware characteristics: speed (not a high priority for many applications), average down time, graphics, memory size, analogue/digital input, hard copy output.
- Nature of applications and problems to be solved: small or large (epidemiological) samples, high or low volume, one-time or repetitive solutions.
- Physical relation of labs and users.

The reader can probably supply additional criteria relevant to his particular environment and applications. In the author's setting, given the type and technical background of investigators, the nature of their problems, he has observed that the user features, i.e., accessibility, reliability, system
From the collection of the Computer History Museum (www.computerhistory.org)

flexibility and versatility, simplicity of operation with no programming, were given highest priority by novice users. As several investigators gained experience and confidence however, their priorities tended to shift toward wanting greater problem-solving power, table-making and data bank capability, visualization of tables and graphs with ability to produce camera-ready hard copy, thereby bypassing manual graphics production costs. As a by-product of observing this growth and developmental process, the author came to appreciate the potential of the appropriately programmed data analysis system as a programmed instruction tool for teaching statistics, graphics and mathematical analysis in the context of ongoing research projects.

It is perhaps appropriate at this point to interject a useful distinction to keep in mind between “state of the art” systems and systems designed to meet specific applicational needs when the primary goal is problem solving. The striving toward “ars gratia artis” seems as appropriate to computer science as it is to the arts and other disciplines. Yet the drive for “state of art” development as an important priority can often be inappropriate to applicational solutions where operational economy, system reliability, accessibility, stability, and ease of operation are critical priorities. Thanks to the “state of art” drive, the computer industry possesses the technical facility and know-how to produce relatively cheap, flexible, reliable data processing systems for identifiable uses. But paradoxically the “state of art” mentality is often associated with an aversion for the so-called “pedestrian” application that has an inhibitive effect on applications problem solving, obscuring the very real value, and even the “state of art” quality in its own right, of applications/problem-solving development.

All of which serves as preface to PROPHET, a biological/chemical data handling system under development by the Division of Research Resources, NIH with Bolt Beranek and Newman, Inc. and First Data Corporation. Its anatomy, physiology and functional characteristics are described elsewhere. What will concern us for the remainder of this article is the PROPHET performance in a specific clinical investigative setting, the Clinical Research Center Core Laboratory of the Harvard Medical Unit, Boston City Hospital.*

FEATURE/FUNCTION APPLICATIONS

Some idea of PROPHET’s applicational relevance to human clinical investigation may be conveyed by describing how specific PROPHET features and functions are utilized in investigative problem-solving, and then how individual investigators have combined various features and functions in PROPHET procedures to achieve specific data processing and analytic capabilities tailored to their needs without recourse to programmers, keypunchers or other middlemen.

The investigative needs of the Unit for the most part require a large variety of statistical and mathematical manipulations, curve fitting, graphing and error analysis on small sample populations, usually on a demand basis. In addition, certain investigators require a capability for construction of open-ended, permanent (i.e., life of project which can be many years and several generations of investigators) data bank which may be easily corrected, extended or updated at any time, and some facility for text introduction and manipulation. Examples of typical projects and their various data processing requirements are given in Table I.

PROPHET meets such needs first and foremost by the “MAKE TABLE” design feature whereby data are entered via keyboard as a permanently stored table which may then be addressed by numbered row or column to perform a variety of statistical, mathematical and graphical operations, using simple English word commands of specified syntax in a dialogue mode. Appropriate commands and dialogue yield graphs (with full control over formatting), straight-line fits to curves, measures of goodness-of-fit to plotted curves, polynomial fits of any degree, correlation studies and so on. The table feature (foreshadowed a decade ago by the stored arrays of the National Bureau of Standards’ OMNITAB System*) has many practical and significant uses and consequences. It forces the investigator to organize himself and his data in a more systematic manner than is often the case, setting the stage for better analysis and therefore better

TABLE I

<table>
<thead>
<tr>
<th>PROJECT TITLE</th>
<th>PROPHET FEATURES/FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolic studies of sepsis and shock in pigs and humans. Liver metabolism studies.</td>
<td>Data bank, statistical and mathematical analysis, graphics, profile analysis</td>
</tr>
<tr>
<td>Studies of parenteral administration in humans and dogs.</td>
<td>Data bank, systems control and accounting procedures, statistical and mathematical analysis, text, graphics, profile analysis</td>
</tr>
<tr>
<td>Substrates and control mechanisms of protein loss in trauma.</td>
<td>Data bank, statistical and mathematical analysis, graphics</td>
</tr>
<tr>
<td>Potassium metabolism and Na-K-ATPase activity in the isolated perfused kidney.</td>
<td>Make table procedures, statistical and mathematical analysis, graphics</td>
</tr>
<tr>
<td>Mineralocorticoid metabolism and relation to O2 consumption.</td>
<td>Multiple-exponential fit and graphics from tables</td>
</tr>
<tr>
<td>Insulin clearance in renal insufficiency.</td>
<td>Trend analyses and graphics</td>
</tr>
<tr>
<td>Effect of exercise or systolic time intervals.</td>
<td>Trend analysis, graphics, statistical significance</td>
</tr>
<tr>
<td>Epidemiological study of cardiopulmonary resuscitation at BCH.</td>
<td>Pen-tablet input of absorption curves for amino acid determinations, graphical and error analyses from tables by column and row</td>
</tr>
<tr>
<td>Relation of amino acid levels to lipid, carbohydrate and protein metabolism in normal and diabetic, pregnant and non-pregnant subjects.</td>
<td>Three-dimensional plotting, determination of non-integer functional dependencies using “Derived” feature of table and linear curve fit commands, correlations</td>
</tr>
</tbody>
</table>

* The Harvard Medical Unit, Boston City Hospital is in process of re-locating to the Beth Israel Hospital, Boston, Massachusetts.

basis for inference. Once the table is made, a camera-ready copy for slides or prints can be made. Statistical or mathematical analysis is accomplished by simply addressing rows or columns, using the appropriate commands (comprising English words and specific syntax) e.g., “Fit line to col 1 vs col 4 of Tablename.” (The PROPHET command language currently accepts only column addressing for statistical and mathematical operations, but several investigators have written row procedures as well.)

Functional relationships of any kind may be computed from any column or combination of columns by the “DERIVED” command in the “MAKE TABLE” dialogue. This feature alone is a significant time saver where large numbers of cases are at stake, yielding derivations that require no number-by-number checking.

As useful as all this is to the clinical investigator in obtaining derived clinical parameters involving such things as drug dosages per unit area, dose response curves, concentrations of labeled drugs corrected for background, and cardiopulmonary parameters from catheterization data, even greater power is realizable in such applications as linear and non-linear regression. Regression is a term statisticians use to refer to a least squares fit of an independent variable against a dependent variable. Most often only one independent variable and a linear dependence (first degree polynomial) is assumed—whence the regrettable but ubiquitous term “linear regression.” Linear regression studies are quickly performed on PROPHET by the command “Fit line to col 3 vs col 24 of Tablename.” The output consists of the best straight line fit, correlation coefficient, standard deviation from regression and several other “goodness of fit” criteria, produced as hard copy in a few seconds.

Appropriate combination of the “Derive column” and “Fit line” features of PROPHET further enable the investigator to free himself from his dependency on “linear dependence” and experiment at will, in a matter of seconds to minutes, with other functional dependencies simply by making columns of any integral or non-integral power of the independent variable, invoking the command “FIT LINE ...” to the appropriate columns and comparing the “goodness of fit” data (which include the correlation coefficient, standard deviation from regression, standard deviations of the slope and intercept, and significance of the slope) for the best fit within the precision limits of the data being analyzed.

What is done for the single variable linear and non-linear functional analyses can be done for multiple variable dependencies—the so-called multiple regression. Here again most clinical investigators who use the technique confine their analytic efforts to linear dependencies, unless other relationships are known to hold. Although multiple regression is available as a Public Procedure only for linear and exponential dependency, PROPHET’s “Derive column” feature combined with the multiple regression feature allow the investigator to test any non-linear functional dependency by a series of simple commands or procedures, opening up regions of functional relationships hitherto rarely explored because of their inaccessibility to the average clinical investigator.

The value of the graph feature resides in the inherent superiority of picture over text for instant and unambiguous identification (to the trained eye) and communication of functional relationships. Thanks to its design, the PROPHET “MAKE GRAPH” feature gives the investigator full control over graph formatting, scaling, titling, curve identification and the like. Because it is quick and easy to use (from data previously stored in tables or arrays) investigators freely use it as it ought to be used—as the first step in determining functional relationships. In the majority of applications visual inspection suggests the most probable functional relationship(s) to be explored, which may then be carried out by making the necessary table alterations and functional dependency calculations. Where several functional dependencies are possible, the “goodness of fit” data may be used as a basis for ruling out one or more possibilities. If all fits yield comparable “goodness of fit” criteria, other criteria, based on additional knowledge of the system under study, if available, may be invoked; failing that, Occam’s Razor (or Law of Parsimony)* is employed.

The “MAKE TEXT” feature finds a variety of uses in footnoting, annotating graphs, and report writing interspersed with graphs and tables.

The pen and tablet entry feature speeds up data corrections or alterations in tables. It also enables the investigator to move graph and table displays to any part of the screen, scaling them at will. Combining these capabilities he may display four to six graphs (depending on size and complexity) simultaneously on one page of copy for sequential analysis studies. The Public Procedure “% Points” allows data entry from a graph or drawing placed upon the tablet, enabling the investigator, via appropriate procedures, to perform curve fits and “area under curve” calculations directly from analogue recorder output tracings.

In all the foregoing applications the hard copier receives steady employment to produce both temporary and permanent copy of tables, graphs, arrays and text displayed on the screen for reference, communication and publication. To the clinical investigator it is an invaluable design feature.

Because clinical investigative data at this center is largely of the small sample type, special attention has been paid to the concept of “sample processing.” Built into the “MAKE SAMPLE” software is the ability to construct a sample from any segment of a table or array, test it for the type of distribution to which it belongs (currently it tests for Gaussian, but will ultimately test for Poisson and binomial as well), and compute central measures, moments (skewness, kurtosis), median, minimum and maximum. Further, the graphics feature encourages the construction of histograms for quick, visual evaluation of population sample distribution characteristics. And the text capability allows carrying along pertinent descriptive explanations and footnotes as the need arises.

*“Entia praeter necessitatem non multiplicanda” which, freely rendered in this context is “If offered a choice between two hypotheses or equations, choose the simpler.” Cf. Mathematical Approach to Physiological Problems by Douglas S. Riggs, Williams and Wilkins, 1963, p. 51.
Why the concern over “sample processing”? Because medical investigators, for historical reasons and the practical difficulties related to extensive calculational efforts (prior to the invention of low cost high speed computing devices) usually assume their data to be normally (Gaussian) distributed; and the mean and standard deviation the only sample measures that possess clinical meaning. Specific exceptions to this generalization are recognized, such as samples of cell counts (which obey Poisson statistics), drug titers (usually requiring log-normal transformations and the use of geometric means) and data that are binomially distributed. Generally speaking however, in all situations not previously established to be other than Gaussian, and in studies of new phenomena, most investigators will assume a normal distribution because few possess the innate curiosity and/or statistical know-how to test for type of distribution. The PROPHET System significantly changes all this by converting a seemingly hopelessly difficult (to the non-statistician) and time consuming task to a simple matter of typing a few words of command and answering “Yes” or “No” as required.

The revolution this obvious and simple design feature portends is suggested by a recent experience with one of the projects passing through the Core Lab data analysis facilities. One clinical investigator became interested in reviewing the status of the white blood cell (WBC) count as a clinical indicator for infectious disease (ID). He arduously collected samples of hundreds of patients in each of several categories: ambulatory patients with and without ID, hospitalized patients with and without ID, and patients with various types of ID. Because “unless you look for something, you’ll overlook it,” he was advised to examine the basic distribution characteristics (beginning with histograms) of several typical samples, though he was strongly inclined by force of habit to accept the work of many investigators before him and assume a normal distribution. To his surprise, he found that all samples possessed skewed normal distributions, calling for use of geometric means and a non-symmetric confidence range rather than the familiar symmetrical confidence limits. With further prodding he saw that the degrees of skewness and the confidence range roughly correlated with the clinical condition, a new diagnostic concept in this setting.

By now excited by the possibilities, he went back to the early studies of WBC counts in normal and sick populations. Some of the early German literature contained tables of raw WBC counts which, on recalculation, yielded results compatible with his own. The entire experience significantly improved his attitude toward statistics and computing systems.

Another example of the relevance of the sample feature to routine clinical investigation occurs in applications of Student’s ‘t’-test, invented by Fisher to provide a criterion for deciding whether two sample distributions belong to the same population sample, or indeed are significantly different. One of the basic requirements of the ‘t’-test is that the distributions be comparable and possess a mean and standard deviation. To fulfill this requirement rigorously, the full distribution characterization (type, mean, variance, sample number, skewness and kurtosis) should be known. If Poisson or binomial, the appropriate central measures must be calculated. If the variances are significantly different (determined by the F-ratio test) a correction is applied, a refinement that becomes important at borderline levels of significance. Finally if the two distributions possess significantly different degrees of skewness or kurtosis, they most likely reflect two significantly different sampling situations, which in itself may have significant clinical or investigative ramifications.

In short, what the PROPHET System sample feature provides for the medical investigator is the statistical expertise and rigor of a Fisher and a Snedecor, with none of the calculational tedium, providing both a learning experience and tool for exhaustive analysis of samples and the phenomena they represent. The investigator may ask more than elementary questions about the functional characteristics and relationships pertinent to his data with full confidence he will be able to get answers almost as readily as he can formulate the questions without calculational tedium, thereby freeing him to think about and test his data in many ways which could result in a stronger basis for inference and prediction.

This catalogue of PROPHET applications to clinical investigation by feature and function is by no means complete or all-inclusive, but it does give some idea of what is possible and provides a framework for describing how these features may be used to solve problems as they arise in the course of a research project, and how they may be combined in procedures, using PL/PROPHET programming language, to produce custom-made data processing systems for specific clinical investigative applications by the investigator himself.

PROJECT APPLICATIONS

At this center, clinical investigative applications of PROPHET fall into two categories—those which accept and utilize its basic design features, utilizing its command language and syntax, and those having additional requirements that are met by recourse to the definition and writing of procedures in PL/PROPHET.

Two ongoing projects involve the construction of a data bank, for which the investigators wrote procedures enabling them to load, delete and alter data in a single command, multiple entry mode, not possible with the usual table commands.

One such study, originating in the Harvard Surgical Unit at Boston City Hospital, involves the collection of some six dozen parameters for a well characterized patient population. Study specifications required the construction of an open-ended data bank allowing indefinite addition of patient parameters with indexing and accounting capabilities, the storing of each patient’s identification, clinical and research data, and the calculation of such output parameters as cumulative weight loss, body surface area, sodium and potassium balance, total nitrogen balance and total nitrogen per unit body area, change in urinary nitrogen, creatinine...
clearance, mean blood pressure, cumulative nitrogen loss, and insulin/glucagon ratio.

The sequence of procedures written in PL/PROPHET to accomplish this, is itself monitored by a system of accounting tables and procedures. A procedure called TRAFFIC, for example, records the name of the table operated on, the name of the procedure called, the date and investigators' initials, the volume of information recorded by type, and the speed of program execution. Two procedures, STATUS and COLCOUNT tell the investigator what kind of data is present in a specified table or set of tables without giving the values. READOUT allows for a quicker display of non-empty values in a series of tables than is allowed by the usual PROPHET table commands.

Output procedures include PRINTNOTES (which prints out a specific patient's study notes), PRINTPTAB (which does the same thing for patient research data tables), RANGESCAN (which prints out the name of any patient specified and the number of values below or above normal of a given clinical laboratory test), GRAPHS (which makes up serial graphs having the same curves and axes from different patient tables) and CODESORT (which searches through a specified set of patient studies and locates a specified set of measurements associated with a specified event code).

In constructing a data processing system to meet the project's unique needs, the investigator utilized the command language and procedures writing capabilities, the text, table and graph features and many of the statistical and mathematical computation functions. Because the design involved many patients, many input and output parameters, with several investigators operating the terminal, additional PROPHET procedures to monitor, oversee and perform accounting so as to stay in control of, and obtain some idea of the magnitude of, the data processing activity seemed appropriate and has proven its worth to the group.

The above project is a typical example of a frequently occurring class of clinical investigative problem—a multi-parameter study designed to permit correlation of a large group of factors with some relevant aspect or facet of the patient's response to a drug, or his overall clinical condition or course, with hopefully, the emergence of some predictive capability as to outcome.

Typical approaches to such multi-parameter correlations with clinical condition are multiple regression; factor, cluster and discriminant analysis; and profile construction coupled with pattern recognition. The latter technique has been utilized by two investigators to follow a group of subjects' clinical courses over a period of time. One such study took 15 clinical observables in a specific order, collected for five series of curves (patterns, signals) that showed the departure of the complete sequence from control (straight line with zero slope, unit intercept) as a function of time, which in this case, correlated with a deterioration in clinical condition (sepsis in pigs). These diagrams clearly identified a group of parameters that on the average remained constant, another that decreased and a third that increased with respect to control as a function of time. It remained for the investigator to take each group and rationalize each correlation in terms of physiologic mechanisms and metabolic pathways. To the best of the author's knowledge, the investigator is still so occupied.

A third frequently occurring application of interest to the clinical investigator is a quick method for computing areas under curves and calculating best-fit functions or non-linear curves. One investigator utilized PROPHET's pentablet input device and the "% Points" entry procedure written by a Bolt Beranek & Newman staff member to determine the area under a Gaussian curve in order to compute amino acid concentration from a strip-chart recording. To utilize this procedure, the analyst simply lays the recording over the tablet and calls the procedure. Then, responding to prompting dialogue he identifies the acid, run number and date, inputs axes reference points, supplies appropriate background levels, enters each point on the tracing by a slight pen pressure on the tablet, proceeding in this manner with each curve until the last. On completion of the tracing, a table identifying each acid peak with its area, half width, maximum point, and associated concentration, together with certain measures of experimental error is printed out. While the PROPHET system is, in principle, ideal for this application, attempts to use the procedure when response time is slow are less efficient than an alternative (but not equivalent) programmable calculator routine.

USER ACCEPTANCE

This account would not be complete without some description of user use, acceptance and impact. By and large as pointed out earlier the average clinical investigator is poorly prepared for mathematical and statistical analysis, and for any technology not directly related to and used frequently in patient care/research applications. Most come to research activity from medical school, internship or residency, steeped in the anatomy, physiology and biochemistry of their special interest career, but somewhat timorous about such basic tools of their trade as bioanalytical instruments, monitoring devices and computers. They recognize the need to "know this stuff," but emulating the example of many senior staff, often try to relegate it to a technician or a knowledgeable medical student spending elective time on a research project.

Accordingly, most users of PROPHET at this location approach it innocent for the most part of previous computing experience, not a little anxious about such "technical" involvement and the threat of competition it represents to "time that should be spent on the wards" or "at the bench." The introduction therefore is one-to-one, personal, systematic, step-by-step, utilizing where possible data from the investigator's ongoing research project. One hour of such instruction usually finds the novice investigator irrevocably hooked. Thereafter most proceed on their own, using the PROPHET manual as a kind of programmed instruction device, directing questions not answered in the manual to other users or to the software systems representatives at Bolt Beranek &
Newman. Most users master enough of PROPHET in this manner to make effective use of it in their research activities without recourse to writing procedures or programs.

CONCLUSION

In summary, the PROPHET System, though still under development, has proved an invaluable research tool in this clinical investigative environment for the following reasons:

(1) With proper orientation it is readily mastered and efficiently utilized by investigators with little or no previous computer experience.

(2) It enables every investigator to apply rigorous statistics and mathematical analysis with a few simple commands, thereby stimulating his development in statistical/mathematical analysis. In this respect it functions as an efficient programmed instruction tool for teaching better biometric design and analysis.

(3) While it provides a relatively wide range of biometrical techniques without programming effort to the user, it also accommodates the user who also can program, allowing him to program and incorporate his own procedures into the basic system.

(4) Its capacity for data bank construction and table sharing has greatly facilitated long-term analysis of data and the communication of data among groups of investigators.

(5) It has been singularly responsive to user problem-solving needs.

ACKNOWLEDGMENT

The author gratefully acknowledges the contributions to this manuscript of the PROPHET users at this center, especially Dr. Robert Auty and Dan Keatinge, arising from the descriptions and documentation of their PROPHET procedures.