Automatic Methods for the Analysis of Physiologic Data

W. J. CARBERY, C. A. STEINBERG, W. E. TOLLES, and A. H. FREIMAN

THE APPROACH of manned satellite flights emphasizes the progress already achieved in rocket and missile technology. Unfortunately there is a concomitant lack of progress in the development of techniques for the rapid analysis and reduction of physiologic data.

When experimental conditions get out of hand in a ground laboratory, there are always direct remedial safeguards available. Therefore the need to perform detailed analysis of physiologic data in real time is not of great critical importance. However, for the man in space, such direct safeguards are impossible. We need to recognize the onset of incipient physiologic trouble at once to provide sufficient lead time to make certain remedial safeguards effective.

The astronaut enclosed in his artificial environment will be at the mercy of any control system malfunction. The sooner we can detect harmful effects of the satellite environment and correct them, the sooner we will be able to decrease the stress of space travel. The extent of useful interaction between the astronaut and the medical personnel at ground control depends on the speed with which intelligent decisions can be made concerning the physiologic status of the astronaut and, in turn, his space environment.

Since the space environment can be simulated to only a limited extent on the ground, the first manned satellite flights must be considered experimental. The high cost of each flight also dictates maximum use of this experimental epoch. Therefore, analytic techniques must be available that permit evaluating each experiment in progress so that alterations in the initial experimental plan can be made during flight. This will increase the amount of new and important information obtained from each flight. The modern high-speed computer can provide for detection of such physiologic and environmental alterations. Incoming data can be rapidly and automatically analyzed, providing medical personnel at ground control with data reduced to a form that can be immediately comprehended and used to make decisions. The high speed and large memory of the modern computer reduces the time-consuming task of detailed data analysis so that critical physiologic data can be analyzed as it is received.

This paper describes a data analysis facility on the ground that can perform sensitive and quantitative analysis in real time of physiologic and environmental data obtained from a manned space satellite. The details of this system are based on the experience

From the Department of Medical and Biological Physics, Airborne Instruments Laboratory, a Division of Cutler-Hammer, Inc., Deer Park, New York and Sloan-Kettering Institute, Memorial Center, New York, New York.

Presented at the Aerospace Medical Association meetings, May 9, 1960, in Miami, Florida.

gained in using computer techniques to analyze physiologic information from clinical surroundings. The problems of analyzing clinical data and space data are fairly analogous and a proposed solution to these problems is discussed, emphasizing the space problem.

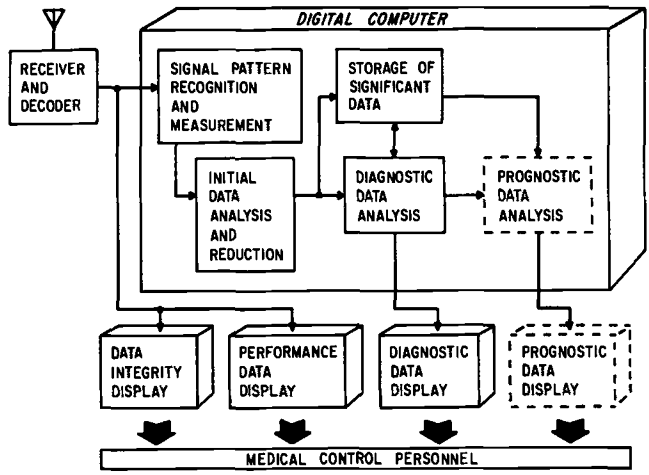

Fig. 1. Major components of ground-based analytical facility.

DATA ANALYSIS

A data analysis facility on the ground should (1) perform a rapid, quantitative, and sensitive analysis of incoming data from the satellite, (2) reduce this data to the significant variables, and (3) provide a composite of this data to use in making decisions as to the well-being of the astronaut and the environment in the satellite.

The tools used to implement this automatic system for rapid analysis of physiologic data are the general-purpose digital computer and associated computer programs and storage media. Other computers, either special-purpose digital computers or analog computers, can be used once the analytic methods are developed. Each analytic operation is statistical in nature and is based on the experience gained with computer analysis of physiologic data.

The major components are shown in a block diagram of the system (Fig. 1). The digital computer and displays present analytic results to the ground-control medical personnel. Within the computer, there are four units that describe the major functions performed by the computer and a unit for storing significant data. Figure 2 shows the sequence of the four types of data analysis to be performed by the computer: (1) Automatic Signal-pattern recognition and measurement. (2) Initial data analysis and reduction. (3) Diagnostic data analysis. (4) Prognostic data analysis. A basic description of the analyses that can be performed in each of these categories follows.

Input Signals.—The telemetered signals are picked up by a receiver and separated into appropriate signal channels by a decoder. These signals can include physiologic, environmental, and performance information. All of these signals would be analyzed in the computer. However, to ensure the integrity of the signals entering the computer, an initial visual inspection of

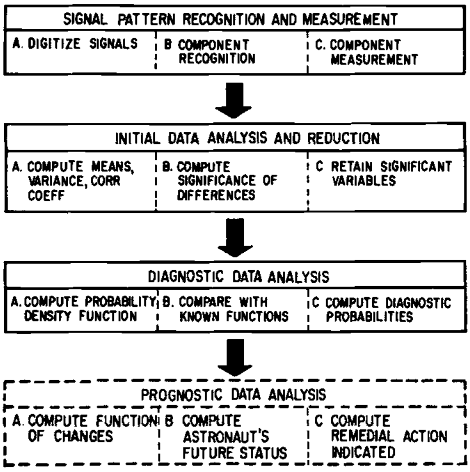

Fig. 2. Sequence of data analysis operations.

these signals would be made by ground-control personnel.

Automatic Signal Pattern Recognition and Measurement.—The incoming signals are classified and measured into (1) signals whose basic amplitudes and times are immediately meaningful, such as temperature, gas tension, and humidity, and (2) signals that have patterns that indicate information of diagnostic interest, such as the electrocardiogram, the electroencephalogram, and other low-frequency phenomena. The components of these low-frequency signals must be recognized and then measured. This recognition and measurement is performed automatically by the computer and the results constitute the input data to be analyzed.

The important feature of this automatic operation is the significant savings in time that manual data measurement and coding would involve.

If the incoming signals are satisfactory after visual inspection, they are converted from analog to digital form—the necessary language for the digital computer. For signals requiring pattern recognition, computer programs based upon time, slope, and amplitude criteria will be used to recognize and measure the components of interest in the signal.

Initial Data Analysis and Reduction.—After the incoming data is measured, it undergoes initial data analysis and reduction. This consists of a sensitive analysis of simultaneously recorded data to detect significant changes that may be taking place in the astronaut and in his space environment. There are at least three factors that contribute to the sensitivity of this analysis:

1. Time simultaneity of data collection. Data from the manned satellite can be simultaneously recorded in time. This means that a considerable amount of the variability associated with data monitored sequentially will be eliminated.

2. Reduction of inter-subject variation. The astronaut population is small and well-defined. Therefore, a considerable amount of detailed information can be derived from preflight test procedures to help set quite narrow limits on the range of normal.

3. Physiologic and environmental interaction. Environmental and physiologic data are available from the manned satellite. Thus, significant changes due to interaction can be sensed earlier than changes in either environmental or physiologic data alone.

The first analytic operation performed on the newly measured data will be to compute the means, variances, and correlation coefficients for a definite time block, for example, ten or thirty seconds of input data. These statistical blocks are then compared to already-stored normal values for each measured variable. The level of significance of any existent differences will also be computed. The variables that prove significantly different, say at the 90 or 95 per cent level of confidence, will be retained for the diagnostic data analysis.

Diagnostic Data Analysis. — After the significant data has been extracted from the total data input, it will undergo diagnostic data analysis. Abnormalities detected in the incoming data will vary in importance, and thus will require detailed analysis before a diagnostic decision can be made with certainty. The goal of the diagnostic data analysis is to determine which of the detected changes are secondary or primary in importance.

The references against which the diagnostic data analysis can be made are a number of abnormal conditions stored in the computer. These conditions will be stored as a series of *n* distributions describing *n* abnormalities. These distributions can be obtained from the results of ground simulation tests on the astronauts or from other clinical experience. Any new set of significant variables coming from the space satellite would be compared to these already stored abnormal conditions; the chance that the astronaut and his environment, or the astronaut and environment together, will show variables consistent with any of these stored abnormalities is thus determined.

The first step in performing this diagnostic data analysis will be to organize the significant data into a form that can be compared with known sets of abnormal data. One such form is the probability-density function. It can, in fact, completely describe the distribution of this new set of data. This probability-density function will be computed, and in turn compared with other already-stored probability-density functions. Probabilities are then computed that will show how closely the new set of variables compares with each of the known abnormal sets already stored. The output of the diagnostic data analysis would consist of probabilities of the form P_A, P_B . . . P_n, where P_A might be the probability that the incoming data is consistent with malfunction of the oxygen supply system, or where P_B might be the probability that the astronaut is undergoing myocardial ischemia due to coronary artery occlusion.

Prognostic Data Analysis. — The goal of this operation is to predict the future status of the environment and the astronaut on the basis of past and present analytic data. The degree of success of this operation is not as certain as the others, but the associated rewards make it worth considering. It is probable that a prognostic data analysis can be performed on environmental data. Physiologic changes induced by the environment should also progress in a fairly predictable manner. However, physiologic changes induced psychologically can be quite erratic and therefore difficult to analyze for prognostic information.

The potential reward of the prognostic data analysis is to provide quantitative information on the urgency of remedial action. Since too sudden a remedial action might prove as harmful to the astronaut as the lack of remedial action, the computer can be instructed to determine a best rate.

A data analysis could provide prognostically useful information in the following manner. Each variable or combination of variables that proves to be significantly different from normal over a given period of time can be expressed as a mathematical function. The computer can be programmed to determine a best mathematical expression for the particular data points available. When this best mathematical expression is determined, the computer can calculate what the state of any variable will be at any given time in the future.

A remedial action could thus determine whether an altered variable has only to be returned to normal or whether additional compensations have to be made in other variables; for example, when the carbon dioxide concentration reaches a dangerous level. In this case, it may be necessary not only to initiate an emergency reduction of the carbon-dioxide level, but also to increase the oxygen level above normal for a short time. The calculation of rate of action will depend on how well the limit of human tolerance to various insults can be determined.

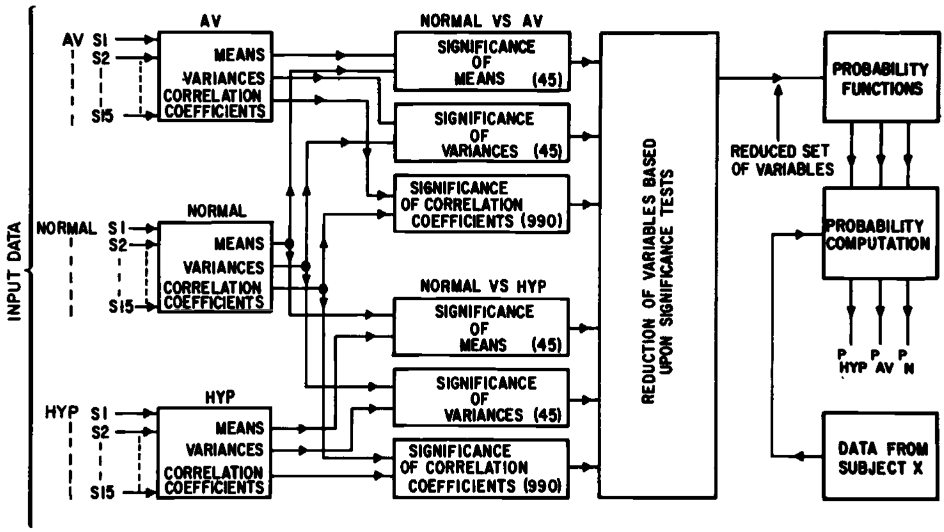

Fig. 3. Diagram of the computational process.

FEASIBILITY OF METHODS

Many of the analytic operations described are analogous to operations performed in the past. The authors are presently engaged in a project, sponsored by the United States Public Health Service,* to determine how computer techniques can be used to assist in the diagnosis of disease, particularly heart disease.

The specific aims of the work performed to date are:

1. To develop a computer method for establishing quantitative classification of normal and heart-disease states.

*This research was supported in whole or in part by the United States Public Health Service under contract No. SAph-70926.

2. To evaluate this method with cardiovascular data from a group of subjects consisting of known normal and pathologic subjects.

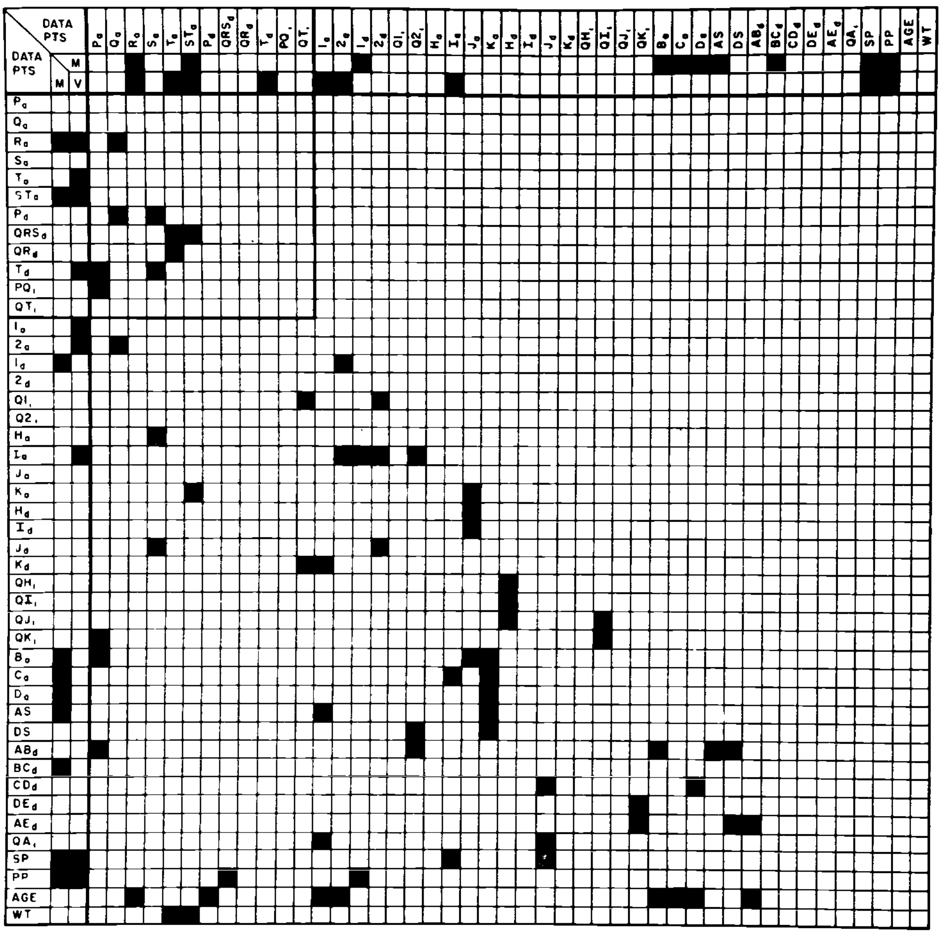

Fig. 4. Significant variables ($\geq$99.9 per cent), hypertensive *versus* normal.

3. To determine the improvement in diagnostic suspicion obtained with comprehensive analyses of several types of cardiovascular information.

Type of Data.—The physiologic signals analyzed were the electrocardiograph, the ballistocardiograph, the phonocardiograph and the arterial pulse—simultaneously recorded in each of forty-five subjects. These subjects consisted of three groups: fifteen normal subjects, fifteen subjects with left ventricular hypertrophy due to hypertension, and fifteen subjects with left ventricular hypertrophy due to aortic valvulitis. In all, forty-five variables, representing time and amplitude parameters, were measured for each subject. This served as the data input for the computer.

Method of Analysis.—An index for comparing each subject with each of the three groups was arrived at by first forming the multidimensional probability distribution for each of these three groups, and then calculating where the variables of each subject fall with respect to each probability distribution. Determining which variables are most significant in this index is accomplish by examining the mean and variance of each variable and the correlation coefficients of each variable with every other variable (Fig. 3). These three quantities completely determine the contribution of each variable to the multidimensional probability distribution. Therefore, the significance of the differences between the mean, the variance, and the correlation coefficient for each variable is an indication of its power in the probability distribution.

Results.—A detailed review of the results of this work cannot be made here. However, certain results do demonstrate the power of the techniques used.

The number of correlation coefficients that proved to be significantly different at greater than the 99.9 per cent level of confidence between the normal and hypertensive groups was 69 (Fig. 4); between the normal and aortic valvulitis groups, 104 correlation coefficients proved significant. Such abundant information is not considered in cardiac diagnosis today, because of the time required to perform the associated mathematical operations manually. In addition, many of these correlation coefficients would not have been available had not each subject's data been recorded and measured simultaneously.

On the basis of the ECG data only, excellent separation was made by the computer between left ventricular hypertrophy due to hypertension and left ventricular hypertrophy due to aortic valvulitis (Fig. 5). This was a surprising result since both groups appeared to have the same degree of left ventricular hypertrophy when only the R-wave was considered. However,

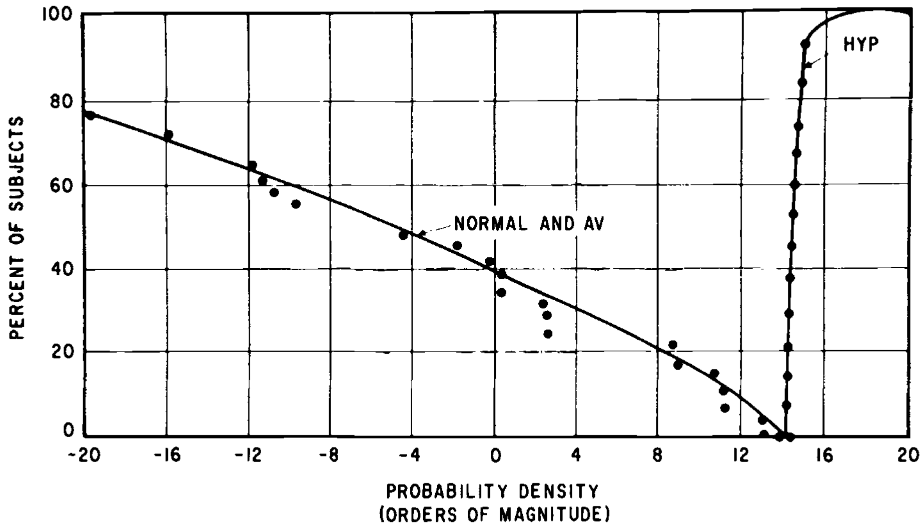

Fig. 5. Distribution of probability densities (hypertensive *versus* others).

when all the ECG variables were considered, clear separation of the two pathologies resulted.

The non-electrocardiographic variables demonstrated differences as significant as those obtained from analysis of ECG variables alone.

Our continuing efforts in this field involve the development of a computer routine for the automatic recognition and measurement of physiologic signals, specifically for the ECG.

SUMMARY

A computer facility for the rapid analysis of data from the manned space satellite has been described. The details of this facility and methods for performing each of the analytic operations are discussed.

The main features of this facility that make possible the rapid and quantitative assessment of the astronaut and his environment are: (1) Automatic signal recognition and measurement. (2) Automatic analysis of several channels of simultaneously recorded data for the early detection of significant changes. (3) Automatic analysis to determine the underlying cause of these changes.

Since this analytic facility does not as yet exist, the results of operating such a facility cannot be presented. However, results obtained from the analyses of similar types of data demonstrate the feasibility of developing such a computer facility for use in experiments with manned space satellites.

Porpoise With A Purpose

The clown of the ocean—the porpoise—has a serious job at Convair in San Diego, where studies of his antics may lead to faster submarines and torpedoes.

"Notty," a one-year-old porpoise, is being studied in a huge tank at the Convair facility. Flow studies are being made as he swims at 20 knots. Temperature sensors are also being used.

Special color dies enable the cameras to follow "Notty"—named after the Naval Ordnance Test Station (NOTS) which is the project manager.

Scientists are searching for the reason why a porpoise has more energy per pound of weight than any other aquatic mammal. It is estimated that a porpoise has ten more power per pound of weight than any other living creature. The answer may bring important answers in under water warfare.

The special NOTS study has been going on for several months and will continue indefinitely. As "Notty" gets older they expect his speed to build up to 30 knots.

Preliminary studies indicate that speed alone cannot be attributed to streamlining. The other factors are what they seek.—From *Western Aviation,* July, 1960.