Requirements for a data processing system for hospital laboratories

by IRWIN R. ETTER
The Mason Clinic
Seattle, Washington

INTRODUCTION

The modern clinical laboratory should employ automation wherever it is applicable in a continuing effort to meet the demand for ever-increasing amounts of analytical information. The laboratory must also keep pace with the latest fruits of medical research which must be added to the clinical diagnostic armamentarium. This laboratory information explosion imposes great demands upon the laboratory staff. Little hope exists for expanding the staff to meet all these demands. Instead, more efficient use must be made of trained personnel. This can best be obtained by automation of clinical test procedures and automation of the accompanying paper work.

Much has been accomplished in the automation of test procedures. Several sets of equipment are now commercially available which relieve the laboratory technician of the tedium of handling each sample at all stages of analysis. Many more will be forthcoming in the near future as instrumentation manufacturers realize the great potential market size and as new instruments become available.

While new automatic instruments perform chemical analyses and relieve the bottleneck at one stage of the laboratory information flow system, if not joined with computers they tend to impose another bottleneck immediately downstream. This is due to the large amounts of time required to perform calculations on the resulting laboratory data and to perform the clerical procedures required by the hospital. At this one facility it has been estimated that the ratio of time spent on paperwork versus true technical work approaches 3:1. The greater the degree of an instrumental automation the larger the ratio.

Typically, the hospital laboratory staff must translate voltage values obtained on an instrument into meaningful values for the clinician, expressed in concentration per unit volume. This alone requires tedious mathematical or graphical manipulation in which gross errors occur due to fatigue. In parallel with the calculation stage, the staff must prepare daily log sheets, laboratory summary and statistical reports, quality control reports and billing information. All of these procedures are clearly within the capability of currently available computing systems. The major difficulty is one of obtaining operating systems at costs that will be economically justifiable. As computer prices drop such systems should become available.

The overall general requirements may be summarized as follows:

1. To minimize routine data processing by laboratory personnel.
2. To reduce error inherent in high volume routine.
3. To provide better utilization of trained technologists for research and development.
4. To reduce data processing time.
5. To provide documented quality control data.
6. To provide predetermined and invariable parameters for acceptable quality of standards, instrumentation drift and deviation of pooled samples, documented during test runs.
7. To provide daily, weekly, and patient summary reports for laboratory, accounting, and clinician uses.
8. To provide these services on-line during the time of actual specimen processing.
9. All of this fully justified economically.

Requirements imposed by instrumentation

Most hospital and/or clinic laboratories have three types of samples to run: routine (hospital), emergency (hospital and clinic) and routine (clinic). Routine hospital samples must be reported out of the laboratory within one day after having been drawn (some specialized tests can take longer). Emergency samples must be reported as soon as possible. This is especially true of hospital samples where test results are a major factor in determining patient treatment. In a clinic, fast response may be required for some diagnoses, but a one-day interval is often sufficient. Ideally, in all cases, only a few hours elapsed time between sampling and reported results should exist. This requires that any data processing system used to process laboratory results should be on-line with the instruments and should be capable of producing results as soon as the sample is run. A laboratory data processing system must interface with two varieties of instruments: automated and manual. Automated instruments operate at a variety of speeds from high frequencies of 1000 cycles/sec. to slow instruments of 1/10 cycle/minute. The speed of manual instruments depends on the rate at which samples can be placed in the instrument and the length of time required for the instrument to reach a steady state reading for the sample. Most manual instruments produce data points at a frequency of about 1/minute. A variety of approaches can be used to monitor the operation of the instrument.

Two basic approaches can be used for the manual instrument. Since an operator is required to be at the instrument while it is running, he can either record the readings and then enter them into the data processing system, or he can signal the system to read the signal on the instrument automatically. Each of these approaches has some advantages. The first approach of recording the data and entering it together with patient number into the system after a run is finished is the least expensive of the two methods considered. A standard keyboard (e.g., teletype or Selectric) typewriter can be used. Since all of the information is entered in digital form, signal processing is not required. Human transference of numbers allows room for error and is the major drawback in the approach. This approach will also be more time-consuming. The more sophisticated approach is to have the computer system joined to the instrument. When a sample is placed in the instrument, it can generate a signal which causes the system to read the instrument signal and search successive readings for a steady state level. This approach is free of error from human data handling and entails less operator time, but may be twice as expensive as the first approach. Further pilot evaluations of the relative economics, accuracy and speed will be the deciding factor between the two approaches.

The critical factors in evaluating the approaches to be taken to handle the automated instrument are frequency content and duration of the signal. Most instruments fall into three groups: High frequency-short duration, medium frequency-medium duration and low frequency-long duration.

High frequency-short duration instruments are typified by high speed mass spectrometers and cell counters. At very high speeds, many data points are generated. Often the scan is repetitive and can be displayed on an oscilloscope. The best way to handle these instruments is to store rapidly and continually the raw data by sampling the signal and converting it at high speed and to process the data at some later time. This is effectively slower than real-time computer operation. An alternative approach is to record the data on high speed magnetic tape for slow replay into the computer. This is probably a less expensive approach since high speed input devices for the computer system can be eliminated. Fidelity of the signal is reduced very slightly by using the tape intermediate, and the costs are certainly reduced. Permanent analog record of the data is also provided in this approach. For some laboratories and instruments this may be an important feature.

Medium frequency-medium duration instruments are typified by the "Auto-analyzer" and gas chromatograph. For instruments in this range, real-time data processing is possible. In this mode, the instruments are monitored about once each second, the raw data being processed as it becomes available. In some situations, short delays in processing may be necessary when exceedingly long routines are required. The effect seen by the laboratory staff has been one of real-time operation.

On-line data processing is highly inefficient for low frequency-long duration instruments typified by amino acid analyzers for which a run may be several hours long, with significant data being developed about once each hour. For these instruments tape recording of data with fast replay into the computer seems to be the best approach. Signal fidelity will be maintained because the signals are of low frequency. Since very few tests run on these long-running instruments are urgently required by the medical staff, the computer can be scheduled to process these tapes during its idle periods (e.g., evenings).

The varied requirements imposed by the laboratory instruments are:

1. Manual entry station for off-line manual instruments.
2. High speed data input channel for on-line high speed instruments.
3. Low speed input channels for medium speed instruments.
4. Analog tape input for fast replay of slow instrument signals and slow replay of fast instrument signals.

Requirements imposed by hospital personnel

Information stored in the data processing system must be readily available to hospital personnel. Laboratory staff, doctors, and research workers all will require the capability of entering information into the system and of interrogating the system as to its contents. These data transfers should require little or no typing skill on the part of the personnel and should yield reports with a minimum of excess information.

The laboratory staff must enter information about test requests into the system. In addition, queries as to the status of on-line instruments and the contents of patient files must be entered. Similarly, information about the test procedures will require modification as procedures are changed in the normal course of laboratory development. Routine reporting of test results on each patient and documentation of laboratory work performed will be initiated by the system itself or they may be requested by laboratory personnel.

The doctors' prime requirement of the system is a supply of up-to-date information about their patients. Test requests are initiated by the doctors who desire clear, concise laboratory reports of the results obtained and their statistical significance with respect to hospital norms. In addition, the possibility of entering significant clinical observations into the patient record should be provided to allow clinical researchers to have as complete a file on each patient as possible. The laboratory data processing system is small and cannot be used to provide a patient record-keeping service. This task should be considered for the hospital information system which is the central information system.

Medical research personnel will require the data processing system to have the capability of being used as a limited capacity, general purpose computer. In addition, those involved in the development of laboratory techniques will require access to data developed in the laboratory in order to evaluate their ideas.

The requirements imposed by the hospital personnel may be summarized as:

1. Flexible information input capability for both fixed and variable format data.
2. Multiple output units with a wide variety of formats.
3. General purpose, as well as special purpose capability.
4. Rapid data retrieval.

Hardware configuration

The description of a system that will satisfy the foregoing requirements is intended as a guide to the perplexed, both the manufacturers and the users of data acquisition systems. Several systems are capable of performing the data acquisition task. Factors of cost and reliability dictate the use of small systems composed of a minimum of devices. Our experience is proposed as a guide to the fundamental elements desirable in a hospital laboratory data acquisition system. At the heart of the data acquisition system is a small computer. The computer acts as both system controller and data reduction center. Several small-to-medium sized machines, currently on the market, have sufficient capability to meet the requirements. The major hazard in selecting a computer is extremism. Too small and/or slow a machine will limit the ultimate capabilities to which the system may be expanded. Too large a machine is a waste of capability. Computer specifications should fall within the following capabilities:

Memory Size	4-16K Core
Arithmetic Unit	Hardware-Add, Substract-less than 3 μ sec.
	Hardware-Multiply-Divide-less than 30 μ sec.

Hardware Index Registers
Real-Time Clock
Flexible Input/Output Structure
- Direct Memory Access Channel
- Data Channels
- Program Controlled Data Transfers
- Program Interrupt

It should be stressed that the instruction repertoire need not be very large since the system will mainly contain one fixed program. It should not, however, require a major effort to modify the program. Reliability of the computer is essential to efficient on-line operation. The computer is potentially the most reliable component of the system and it should be expected that computer failure be a rare occurrence.

Peripheral devices (i.e., tape recorders, typewriters, printers) have been of notoriously low reliability. This dictates the use of minimal number of peripherals and wherever possible some duplication. The system should be designed to allow operation to proceed despite any peripheral failure.

In any data acquisition system of more than minimum size, some mass storage device is necessary. Exceedingly high speed access is not of major importance. Reasonably long times (up to 1 sec.) can be tolerated to retrieve a record from the mass storage device. The majority of this time would be spent in finding the location of the record in the mass storage and only a short time spent on actually performing the transfer. The processing system must be capable of sending a request to the mass storage device and then continuing operation until the actual data transfer takes place. Reliability is of great importance here since system operation would be drastically curtailed if the mass memory unit became inoperable.

The basic communication channel between the system and the hospital personnel is through the keyboard typewriter/printer unit. The most widely used unit is the teletype. At least two are highly recommended and the most rugged versions should always be used. One unit should be equipped with a paper tape punch and reader. Only in very large systems will a high speed line printer be worthwhile.

A slow speed card reader should also be included. Information entered on cards includes: test type information, to be entered only when modified, and any standard input information to relieve the need for typing. It is expected that the card reader will occasionally not be functioning properly.

The analog and digital interface which allows direct monitoring of laboratory instruments is composed of scanners or multiplexers, A to D converters, and a set of digital logic for testing instrument operating conditions. Two analog input systems can be used for fast and slow input respectively. The slow speed unit should be capable of accessing 200 points/sec. while the high speed unit should be capable of at least 10,000 points/sec. Both units should operate under program control. The digital interface should be able to detect voltage above a preset level and switch closures, generating logic levels 1 and 0 for true and false states respectively. In addition, the capability of reading a set of voltages or switch closures in groups equal to computer word size and organizing them into computer words is essential.

The signals on most slow and medium instruments can be taken from the swinger of a retransmitting slide-wire mounted on the recorder. The swinger is connected to the input of the scanner. The scanner in the slow speed input unit is sequentially connected to each position. At each position are several signals: 1) a ground reference, 2) the signal from the swinger, and 3) any logic that may be used to test instrument conditions. From the scanner, the analog signal is sent to the analog-to-digital converter which is under program control.

The digital interface consists of a battery of logic elements which interrupt the computer when instruments need to be serviced. Instrument priorty may be established either by hardware logic (e.g., a daisy chain in which instruments connected closer to the system have higher priority) or by software routines which check a sequence of interrupt sources; the sequence establishing the priority. Also included in the digital interface are control units located at the instrument. These units are used by the laboratory personnel to inform the system of the status of the instruments at that location. In addition, any manual data entry stations are processed by a digital interface which essentially reads a special purpose keyboard.

Software configuration

The key element in obtaining a working data acquisition system is the software. Many shortcomings in the hardware can be offset by creative programming. There are basically four parts of software: executive routines, data processing routines, bookkeeping routines, and communication routines.

The executive routines are responsible for controlling the operation of the system. The order and timing for servicing interrupts are its most important responsibilities. Many different approaches to priorities can be used. One simple approach is to perform a minimal amount of processing on each interrupt as it occurs. The interrupts which are waiting and have not been completely serviced are then checked to determine which is oldest and the oldest is then processed. An alternative is to develop some method of deciding how much is to be done with each interrupt and handle the shorter ones before the longer ones. Another scheme which has merit is to assign levels of processing required of each interrupt, e.g., level 1 will be storage of raw data, level 2 will be data smoothing, level 3 peak picking, etc. Level 1 is performed on all interrupts. Then, if time permits before the next interrupt, any level 2 processing jobs will be done. Certain levels like level 3 will result in assigning higher levels to the data (e.g., performance of standardization or conversion of a peak value to concentration units) or will result in completed processing of that piece of data. It is assumed in all these priority schemes that there is sufficient time to process all the interrupts since there will be slow periods when less work is required by the system. The executive routines also control the sequencing of the printout routines, assigning highest priority to those messages which require some action on the part of the laboratory staff (e.g., a warning that the standard samples are not in the proper ranges) and lowest priorty to printing long summary reports.

The data processing routines are called in a variety of sequences by the executive routine to perform all mathematical manipulations of the data stored in the computer. For most on-line instruments, the data processing steps of "smoothing" and "peak-picking" are common to several types of instruments, while others such as the calibration and unknown determination routines may be specific to one type of instrument. Development of these data processing routines requires a thorough understanding of the laboratory methods. Also, included in the data processing routines are a group of subroutines that are used by many different routines. Such subroutines as floating point addition, multiplication and division, binary to Binary Coded Decimal (BCD) and BCD to binary conversion routines, transcendental functions, and various logic routines that are not performed by hardware registers are included in the set.

Bookkeeping routines perform all the data sorting required by the system. This includes sorting test requests into lists of samples to be run on each test type, relating calculated data on an instrument run to the patient's record where they are to be filed, and assembling data in formats required for typeout routines. No data manipulation is performed by these routines except for data conversion between the coding required by the input/output peripherals and the binary coding of the computer.

Communication routines perform all the operations of acquiring data into the system from the instruments and the peripherals, and developing reports for the laboratory and hospital staffs. These routines include the testing of instrument status (on, off, test type, samples expected, etc.), obtaining analog data from the instruments in the digital form required by the computer, obtaining information entered through the teletypes and card reader, printing of laboratory work reports (instrument load tests, daily work reports, billing reports), and patient record reports on a test request basis and on a summary report basis.

In the normal mode of operation, the executive routine resides in core and calls routines as needed from bulk storage into core. Sections of core are also reserved as transfer stages between core and bulk storage and between core and peripheral devices. In addition, all data of current interest to the system are stored in core including fixed information such as values of test standards and the information currently being developed by the instruments. The executive routines also reside in bulk storage to eliminate the need of loading it through the paper tape reader. Initially, a small paper tape routine is entered into the computer which when executed enters the executive routines into core and starts the operation. The executive routine need not be entered into core again unless it is determined that part of the routine has been lost.

All software routines are permanently stored on paper tape and cards. In addition to the full system's normal routines, several operable routines of a more limited scope are available to avoid using particular components of the computer should any part of the system be out of operation. Thus, if the bulk storage is inoperable, a routine can be used which merely processes the data as it is developed on the instrument and produces reports based on laboratory runs. Data sorting according to patient is eliminated as is summary reporting. All data produced during this period are stored on paper tape so they can later be entered into patient records when the bulk memory is again operable.

System development

Since the system described here is large and the operation complex, it is recommended that development be done stepwise. This will allow the laboratory staff to gain confidence in the system's capability and not frighten them with a large "monster" they cannot understand. Laboratory personnel should be closely involved in the development to aid in obtaining acceptance of the system. The cost of such a system is relatively high for most hospital laboratories and the stepwise acquisition of the system greatly eases the strain on the budget. Almost all that is done during the early stages will be of use at later stages.

The first stage is the development of a unit which will acquire data from several instruments of the medium speed variety in an on-line, real-time mode. No patient identification or sorting is performed by the system at this stage. All results are in the form of a report on an instrument run. (Correlation between position on the run with patient is still done by hand.) The report, developed at this stage, will be used in later stages as a laboratory work report to provide documentation of laboratory procedures. At present pricing levels, cost of this initial system can range between \$35,000 and \$100,000 depending on the size of the computer and the size and type of interface (instrument and human) developed.

This level is a convenient point to use on the first stage since it will not appreciably disturb the information flow in the laboratory and will relieve the laboratory staff from hand mathematical processing of instrument data. Since little peripheral equipment is involved (teletype, analog, and digital interface) reliability of the system can now be proven. Laboratory staff should become involved with the system to establish complete confidence in the system's capability.

Further development is aimed in two directions, increasing the number of instruments tied into the system and increasing the amount of the paper work load assumed by the computer. As new instruments are put on-line with the system, thorough testing of the data processing must be performed. Addition of peripherals such as a card reader will enable the computer to assume the job of preparing load lists for the instruments. Bulk storage will allow for the keeping of patient records and reporting of results on a patient-by-patient basis thus eliminating two of the major paper work jobs now performed by laboratory personnel.

The major guidelines at all times should be proven capability, system reliability, economic justification for expansion, and maintenance of system flexibility. At present a reasonably sized data acquisition system for a hospital laboratory will be capable of accepting test requests on punched cards, providing instructions to laboratory personnel for the running of instruments, automatic retrieval of data from about twenty automatic instruments and several manual instruments. It will also provide documentation for the laboratory, the medical staff, and the billing office in terms of test results, individual and summary reports on each patient, and amounts to be charged to patient accounts. While little or no attention to the inner workings of the system should be required during normal operation, in-house capability to modify the computer system will become desirable if the unit is to keep pace with laboratory developments. At present price levels in the computer industry, such a system as described here will cost approximately $250,000. This is more than the market (with a few exceptions) will be able to afford. A price tag closer to $150,000 will be more acceptable for a system that is completely assembled with a complete set of operating software.

As hospital computer systems of several varieties are developed and become accepted pieces of equipment, an integrated hospital system will be required. Each separate unit such as the laboratory data acquisition system should then act as a satellite unit exchanging information with a central hospital information system. The need for a separate laboratory unit will then become especially great. The central hospital system will be occupied with retrieving, assorting, and supplying information to and from all parts of the hospital and should not be required to devote large amounts of time to processing laboratory information. With a laboratory data acquisition system preprocessing the laboratory data into a form that is compatible with a central information system, the laboratory will be able to make highly efficient use of the central system.