The London Hospital

Computer Facilities at The London Hospital

COMPUTER FACILITIES AT THE LONDON HOSPITAL

1. General

If any project is likely to produce a substantial quantity of data for analysis, or if some extensive calculations or tabulations are required, it is desirable to consider the possibility of using the hospital's computer at an early stage in the project - certainly before detailed work is started or the design of survey forms is finalised. Without this sort of preparation it is often difficult at a later stage to obtain the full benefits of the available computer facilities or else a great deal of unnecessary data transcription and editing is required.

2. The Computer Facilities

The hospital's computer is an Elliott 803 and it is capable of storing just over 8000 12 digit numbers simultaneously. Five hole paper tape is used to transfer information into and out of the computer which has two paper tape readers, two paper tape punches and a teleprinter. The backing store consists of magnetic film; a single reel of which will hold about $\frac{1}{4}$ million 12 digit numbers or 10 million bits. Although there are problems for which it is in some way inadequate, it is likely that this computer will be of considerable assistance to most projects. In spite of the growth of use of the computer (the average running time last year was about 105 hours per week) some staff are still not aware of the considerable power of this machine. The bulk of the effort of the Computer and Operational Research Units is directed to the design and implementation of routine computer systems to assist the functioning of the hospital; and these systems make substantial demands on the available resources. By contrast research projects can be handled much more simply since they can be fitted into the punch-room and computer schedules at convenient times. If programming is required, it is possible for many staff to handle this themselves with a minimum of advice and assistance.

3. The Approach

Although a person may be able to conduct his research on a fairly vague "do it and see" basis, this is only possible because he can cut off unpromising lines fairly quickly and alter his plans to take advantage of results as they appear. The current London Hospital computer is less flexible and it is necessary to specify the objectives to the following extent:-

(a) Input data in order, type, range and where necessary including form design.

(b) Internal arithmetic and logical operations making allowance for all combinations of input and derived data.

(c) Results to be output in order and extent.

Thus, until the research worker has achieved at least some clarity with regard to the manipulation of his data, there is very little point in involving a computer. However, as soon as the preliminary hypotheses have been formulated to the extent of indicating that certain statistical tests are required, or a particular formula is to be tabulated, or certain types of data are to be stored for future reference, it is then possible to examine the computer content of the project. It should be emphasized that it is rare for a project to be specified precisely and for the requested results to provide all the information required. The preliminary results invariably suggest the need for further investigations but the clearer the matter is seen initially the easier it is to make appropriate allowance for re-examination of the data. A radical change of plan can often render useless months of programming effort and hours of computer time. It should be noted that one or more test runs on data for which the results are known are normally required and it is normally necessary to ensure that the test data includes all the variations encountered in the working data. Although this stage often involves much tedious checking and calculation, the reliability of the final results is directly dependent upon the thoroughness of this testing process.

As indicated above, much research involves a dialogue between the research worker and the computer. This is necessary because the research worker must know precisely what results he has got and make sure that they are exactly what he requires. For instance, if a file of data on patients is searched for those over the age of 15, a search will obviously exclude patients whose age may have been unknown when the data were collected and had the age entered as zero. The dialogue between the research worker and the computer is terminated when the former is satisfied that he has the results he needs.

4. Detail

Having clarified the outline of the problem, there are three ways of proceeding:-

(a) Standard programs from the program library may be used.

(b) Sufficient programming technique may be acquired to enable the research worker to write his own special purpose programs.

(c) The Computer Executive may be convinced that the problem is of sufficient importance to the hospital to allocate it a high enough priority to enable the Computer or Operational Research Units to take it on as a project.

In practice, very few extra projects can be undertaken by the computer staff and thus it is important to know what is available in the Program Library before embarking on a crash course in programming. Quite apart from the manufacturers library there is an extensive library of programs available from other users of the Elliott 803 computer. In addition many programs written in Algol (an international scientific programming language capable of being used on many machines) may be modified to run on the computer. No listing of these resources is feasible in a short report but the program descriptions may be inspected in the Computer or Operational Research Units. In general one may expect to obtain library programs to execute certain well-defined calculations - e.g. standard statistical analyses. A selection of the more useful programs is given in the appendix. The best method of proceeding with the project can usually be settled once the volume of data, the variety of analyses and the time scale of the research project are specified.

5. Research

One only obtains useful results from computers by taking great care with the data input, with the programming if such is needed, and with the interpretation and examinations of the results. Poorly prepared data combined with the inadequate understanding of the program and results can combine to produce dangerously misleading results. The computer staff are ready to discuss projects and assist in helping others to learn the skills of computer operating and programming but in the present circumstances this mainly is an advisory service not a complete programming service.

The computer facilities are there, it only remains for as many staff as possible to examine them in the light of their own work to obtain the greatest benefit for the hospital as a whole.

Appendix

A. General Survey Programs. These programs are still being added to and modified to provide extra facilities. The basic programs currently available are:-

(a) The Data Input accepts data in a standard form, sorts into a particular order and stores it on magnetic film in a compact form. Each record can consist of up to 2 separate groups of alphabetical information such as a patients name (each group must be less than 24 characters), 15 twelve digit integers, 36 small positive integers (each less than 8191) and 1064 answers to numbered questions having the answer yes or no. This program has been used for storing diverse information from the specialist records of certain patients (e.g. obstetric file) to data about the emergency anaesthetic

service. It is the result of considerable experimentation and it has been designed for simple, economical form entry. Many research file creation projects can be adapted to use this program without much trouble.

(b) General Summation Program will read data stored by input program, unpacks it into a convenient form and provides a simple summary. Modifications can easily be made to the basic program to carry out more complicated calculations on the records.

(c) The Fact Listing Program examines all the stored records and provides a list for each of the numbered questions of the records for which this question has been answered with a yes.

(d) The Search Program enables particular records to be extracted.

B. Multiple Regression (LS17) gives simple basic information about the variables input, provides a correlation matrix and enables either manual or automatic regression to be carried out.

C. General Statistics Program provides simple information about samples such as the mean, standard deviation, standard error of the mean, confidence limits as well as enabling t-tests to be carried out to test for difference between sample means.

In addition, there are readily available programs for handling statistical techniques such as analysis of variance, principal component analysis, auto and cross-correlation analysis.