Computer processing of biochemical information without going on-line

I. D. P. WOOTTON

From the Royal Postgraduate Medical School of London

It is now abundantly clear that the assistance of a computer is of great value in the day-to-day work of a large hospital biochemical laboratory. If immediate access to the computer is required at any moment, then either the computer must be reserved for the exclusive use of the laboratory staff or else an extensive central installation must be available with the ability to accept multiple inputs on a time-sharing basis, one input terminal being located in the laboratory. Both these approaches are expensive. In the first case, the laboratory must bear the whole cost of the equipment and for practical reasons this may result in the provision of a computer too small adequately to do the rather complex tasks involved. On the other hand, although a large central computer may provide relatively cheap processing power, it is no small task so to design the central system that the laboratory can have unrestricted access at intervals of a few seconds throughout the working day, *ie*, whenever instrument readings must be recorded or the laboratory records are interrogated. It is likely, therefore, that off-line computers will continue to be used where the input data are collected in the laboratory and sent once or several times daily to a computer for batch processing. This does not even imply that the computer must be within physical reach of the laboratory, since a data link, *eg*, by telephone wire, may be possible.

Of course batch processing laboratory results imposes certain limitations compared with an on-line connexion. These problems have been faced during the last two years in the biochemical laboratories of the Royal Postgraduate Medical School, during which time we have had access to a medium-sized general purpose digital computer. The purpose of this paper is to indicate the areas in which we have found the computer to be valuable and how the disadvantages of off-line working can be minimized.

THE INSTALLATION

The equipment was commissioned in 1967 and includes a NCR/Elliott 4120 central processor with 16K of 24 bit words in core storage and four magnetic tape

handlers. The configuration is designed for use with paper tape input and output and includes two paper tape readers and two punches. The lack of a fast output device such as a line-printer has proved to be a disadvantage and all output is printed up off-line at a rate of 10 characters/second.

MATHEMATICAL APPLICATIONS

AUTOANALYZER CALCULATIONS A large modern biochemical laboratory usually deploys a number of Auto-Analyzer channels, often grouped together into multi-channel instruments. As each channel is capable of producing 40 to 60 peak values per hour, the task of converting these to concentration values consumes a considerable amount of skilled effort. Many laboratories have described methods whereby computers can assist (Gould, 1967). This has usually involved some method of 'peak picking' by modifications to the recorder (Flynn, Piper, and Roberts, 1966) or else by taking repeated readings of the instantaneous value of the channel every second or two and finding the peak by comparing successive values (Blaivas and Mencz, 1966). Because a very large number of values is involved, the latter technique is only possible with on-line instruments.

A third method of data acquisition is to use a 'trace reader,' which comprises a flat bed to which the Auto-Analyzer charts can be clipped. An electric probe is then used to pierce the chart paper at the position of each peak in turn. A paper tape punch connected to the trace reader operates each time that the probe is used and a 'word' of data is punched into the paper tape. The word includes characters which describe the channel under examination, the specimen number, and the peak height, that is, the height above the baseline of the chart.

When the paper tape is subsequently fed into the computer, the program compares the peak heights of the unknown samples with those of the standards, interpolating between the nearest pair of standards and converting the peak height into the required concentration. The AutoAnalyzer plates are always loaded in a standard manner and premarked plates are convenient. In addition to the standards and unknown samples, 'drift' samples are incorporated in selected places; the program calculates the apparent value of the drift sample each time that it is encountered (Fig. 1). An unacceptable rate of drift leads to a halt in the calculation and an error message. If the drift is within tolerable limits, the values

FIG. 1. AutoAnalyzer chart ready for the trace-reader. The drift line established by the drift samples has been marked, also the 'pool' serum used as a laboratory control.

of the unknowns are corrected by an amount calculated to be appropriate for the particular position of the peak in the run.

Provision is also made for cancelling wrongly punched peaks, and for inhibiting the calculation when peaks are not in range or are otherwise faulty. For specimens that are prediluted, there is a facility to include the dilution factor in the arithmetic.

The output of the computer is in the form of a list of results, prefixed by a specimen number, and with any quality control specimens marked for easy reference. This method of reproducing the results is especially valuable in the case of multichannel analyses and provides the results in a form well suited for preparing reports and filing either manually or by computer (Fig. 2).

QUALITY CONTROL CALCULATIONS The mean of the results in any batch of analyses has been shown to be remarkably constant, provided that (a) the batch is large enough and (b) that outlying results are truncated in certain instances where the abnormal results are numerically very different from the normal, *eg*, in the case of blood or serum urea (Hoffman and Ward, 1965). Calculating the daily batch mean and examining it for consistency has become an important method of quality control. It is readily carried out as an addendum to the AutoAnalyzer peak calculating program.

A more sensitive use of the daily batch mean value has been described by Whitehead (1966). A cumulative sum ('cusum') is calculated each day by adding in the

difference between the daily batch mean and a formal reference value. This technique, which is widely used in industry (Woodward and Goldsmith, 1964), is normally applied by manual graphical plotting and inspection. In place of this, the computer analyses the most recent part of the graph and tests the consistency of the results. It prints out an abbreviated message applying to each channel which states whether the method is in control or not. These condensed results are displayed in the laboratory for the information of the laboratory staff; in the event of a method running out of control, the computer on demand will print out the details of the calculation. This helps to pinpoint the period when the trouble began and enables effective action to be taken.

HANDLING LABORATORY RECORDS

In this group of applications, the computer is not concerned with mathematical calculations but is used to store and sort information.

THE SYSTEM We assume that the following are desirable properties of a laboratory record system.

The results of all tests are stored. Each result is linked to a single patient so that it is possible to record all the laboratory tests done on any individual, although for technical reasons of storage space, the older results may be filed away in less accessible form.

On each occasion when a test is done on a patient, the more recent results (at least) of previous tests done on the same patient are retrieved, so that the laboratory staff

= 24 1 69					
ELU					
3208	132	4.3	94	27	79
3212	136	3.6	93	30	20
3235	137	3.8	99	28	22
3247	139	4.0	101	29	35
3255	133	3.8	96	28	29
3099	133	4.5	93	24	157
3269	139	4.0	102	26	164
3275	NN	NN	NN	US	77
3276	138	3.4	97	29	56
3280	133	4.4	99	26	40
3283	133	3.9	99	26	51
3285	138	3.7	102	25	29
3291	132	4.4	99	23	101
3292	136	6.0	102	24	111
3295	134	3.1	100	25	21
3296	141	5.1	100	33	66
3297	134	4.2	97	29	39
3298	139	4.2	95	33	46
3310	140	4.0	99	28	29
3311	145	5.3	105	27	57
3312	138	3.8	98	27	47
3314	142	4.1	103	31	42
3315	141	3.6	99	32	41
3316	137	4.1	98	30	46
3317	141	3.4	103	30	39
3318	141	3.8	105	24	34
3319	141	4.3	104	28	42
3322	138	3.6	106	23	21
3323	137	3.8	105	23	20
3324	134	4.1	103	22	19
3327	126	4.2	90	28	48
3328	NN	NN	NN	NN	NN
= POOL 140 4.6 100 21 55					
= POOL 141 4.6 101 22 57					

FIG. 2. Computer-calculated results from a five-channel (electrolyte and urea) AutoAnalyzer. The columns in order are specimen number, sodium, potassium, chloride, bicarbonate, and urea. NN represents a peak of unacceptable quality which must be repeated; US is printed when the sample is too old for analysis. The pool (control) results are printed separately.

have the opportunity of comparing the new result with the previous ones (Wootton, 1965).

Cumulative reports are prepared for the wards, comparing new results with the previous ones, and displaying the results in an integrated and compact form (Flynn and Vernon, 1965; Whitby and Owen, 1965).

Following the flow of information through the laboratory, the first task is to prepare a 'day list' (Fig. 3). This is punched off-line as a paper tape and provides a record of the requests received by the laboratory. The items punched include the basic identification data relating to the patient (name and unit number), the accession number of the specimen, and the analyses requested. By

input to the computer, the day list is 'exploded' into individual test requests and sorted into separate work sheets (Fig. 4), each sheet representing the work required of an individual analyst or given apparatus. Blank spaces are provided on the sheets so that the results can be filled in; if AutoAnalyzers are used, the output of the calculations is already in the correct form.

Thus far, the system has been in daily use for some time and is now indeed an essential part of our routine. The remaining part of the system has worked in experimental trials, but is not yet done routinely. This begins with punching the completed work sheets, which are then fed into the computer. The computer sorts the data, re-

= THURSDAY JANUARY 23RD 1969 PM (1)						
*23/01/69						
HUGHES	348555	MU	3131	S	*SERIAL URINES 5*	
RIDGE	1876	EB	3132	B	ELU UA TP BIL:	
JACKSON	348350	OP	3133	B	ELU:	
EAGLE	341737	OP	3134	B	ELU:	
ROBERTS	96525	G3	3135	B	ELU LDH AG AP ENZ:	
ANGELINAS						
ANGELINETTA	276042	I3	3136	B	ELU CA AP BIL:	
SHEVLANE	97738	I3	3137	B	ELU:	
BUTLER	218086	C6	3138	B	ELU LDH BON :	
HEWER	348440	D1	3140	B	ELU AP BIL:	
HOPKINS	86191	D1	3141	B	ELU BON AG:	
GREGORY	347132	D2	3142	B	UA:	
CURTIS	347132	D2	3143	B	ELU:	
DELAPENHIA	343764	D4	3144	B	AMY ICD:	
PRICE	346718	B6	3145	B	FBI:	
HOWITT	148861	B6	3146	S	*IMMUNELECTROPHORESIS*	
PALLATT	338760	B4	3147	S	*GTT*:	
TALBOT	190119	OP	3149	B	ELU:	
BURT	336890	OP	3148	B	LDH LFT:	
KARANJA	349433	OP	3149	\$		
KARANJA	1877	OP	3150	B	AG:	
VOLLER	198286	OP	3151	B	ELU:	
SIMPSON	348514	OP	3152	B	ELU AG:	
EMONS	336402	CA	3153	B	LFT ENZ AMY:	
JEE	131837	D2	3154	B	ELU BON LFT:	
HASTINGS	301806	G5	3155	B	LDH AP BSP ENZ:	
DWYER	337586	G5	3156	B	CA P TP:	

FIG. 3. A portion of a day list for computer input. The requests are coded: eg, BON represents calcium, phosphorus, and alkaline phosphatase; AG represents total protein, albumin, and globulin. A\$ cancels the line in the case of a mispunching.

assembling the various items from different work sheets relating to each patient. It then reads down the main laboratory results file which is held on magnetic tape in patient order, interpolates the new results in the correct place, and rewrites the enlarged file back on to magnetic tape. At the same time, any patient record which has been altered during the updating process is also copied on to a spare magnetic tape so that this tape can be used to prepare the day's reports. These may be the results of a single test, cumulative patient reports, or ward reports, in which are summarized all the tests just performed on the patients in a single ward.

SEQUENCE OF OPERATIONS For reasons of convenience, we have kept the number of daily computer runs to a minimum and have now concentrated these into two each day. They are at fixed times related to the rest of the laboratory activities. Specimens (such as blood sera for electrolyte which are analysed daily) are accepted for reporting on the same day if they arrive in the laboratory before noon. As soon as they are received, they are day listed and the first computer run of the day at noon produces that day's work sheets. At the same time, trace-reader data tapes relating to the specimens arriving during the previous afternoon and during the early part of the morning are computed, and quality control calculations are performed.

During the second run, about four hours later, the remaining data tapes of the morning's specimens are processed, so that all these results are available for reporting on the day that the specimen was taken. This is also the run when it is proposed that manually completed work sheets will be input and computer-prepared reports will be available.

Each run occupies about 15 minutes of computer time. The equipment has been very reliable and we have only rarely been unable to complete our routine runs. However, complete record keeping and reporting requires, at the moment, correct functioning of all the peripheral equipment, including all four magnetic tape handlers. These are less robust and reliable than the computer. This is the principal reason why we have not yet implemented the complete system.

URGENT REPORTS AND ENQUIRIES The method of working described is essentially dependent on a smooth and organized flow of operations. It can be completely disrupted by attempts to insert urgent specimens. For this reason, we have established an adjacent 'hot' laboratory, largely using manual methods, to meet urgent needs. Results from this laboratory are merged later with the main bulk.

Urgent enquiries cannot be answered from computer-stored information in this system except occasionally

PAGE 1	RUN	24/01/69				
3-CHANNEL WORKSHEET						
		ALB	GLO	TP	P	AP
23/ 1/69						
RIDGE	3132	XX	XX		XX	XX
ROBERTS	3135				XX	
ANGELINETTA	3136	XX	XX	XX	XX	
BUTLER	3138	XX	XX			
HEWER	3140	XX	XX	XX	XX	
HOPKINS	3141					
BURT	3148				XX	
KARANJA	3150				XX	XX
SIMPSON	3152				XX	XX
EMONS	3153				XX	
JEE	3154					
HASTINGS	3155	XX	XX	XX	XX	
DWYER	3156	XX	XX			XX
RAYMOND	3162				XX	XX
CHARLES	3163				XX	
RODRIGUEZ	3164				XX	
WILLIAMS	3165	XX	XX			
SIMARD	3171	XX	XX	XX	XX	
MCCORMICK	3172					
KENT	3185	XX	XX			
WHELAN	3188				XX	XX
AYTON	3189	XX	XX		XX	
NEALE	3192				XX	XX
DALLY	3193				XX	XX

FIG. 4. A worksheet for the AutoAnalyzer doing proteins, phosphorus, and alkaline phosphatase. The patients' names and specimen numbers are listed. If any estimation is not requested, the computer fills in the corresponding space with XX, thus the analyst is required to complete the vacant positions.

if they are received when the computer is actually performing a laboratory run. We have found it convenient to file in order the day lists, completed work sheets, and computer-calculated results, which together constitute an ordered record of the total work. Although not perhaps ideal as a manual file of results, the assembly is adequate for occasional use to answer telephone queries.

OTHER APPLICATIONS With regular access to a computer, and results in computer-readable form, a number of other applications become possible. These will vary from laboratory to laboratory, and even from time to time in a single laboratory, so no attempt is made to provide an exhaustive list. We have found the following useful: a statistical package program which is in almost daily operation for performing significance tests on experimental data (Wootton, 1968); programs to construct the frequency distributions and calculate population statistics of laboratory results; a program to compare laboratory day lists one with another and to provide information on the correlation of one set of requests with the other; a program to mark laboratory results with defined characteristics, *eg*, values higher or lower than preset limits.

CONCLUSIONS

Although the introduction of computer methods into a routine biochemical laboratory has not been without its problems, the advantages gained are certainly very great and enable us to do things which we otherwise could not contemplate. The laboratory staff have not found the adaptation to the computer methods any more difficult than any other innovation of a comparable importance and welcome the ability of the computer to free them from certain dull routine tasks. Staff establishment changes have been minimal; all the laboratory punching of data and computer operations can be handled by a single operator. Program maintenance is a part-time responsibility of a single programmer, assisted by several members of the laboratory staff who have added program writing to their other accomplishments.

For several years to come, it is likely that off-line processing on computers will meet the needs of many laboratories; certainly such systems are much more easily introduced than are on-line ones.

SUMMARY

The Biochemical Laboratory of the Royal Postgraduate Medical School has had two years' experience of the use of an off-line digital computer for the calculation and handling of biochemical results. This method has proved to have certain advantages over the alternative on-line working. The laboratory working programme is carefully scheduled so that access to the computer is only required twice daily. Delay in the production of results is minimized and the flow of work proceeds smoothly. The facilities which the computer offers include listing of requests, calculation of AutoAnalyzer results, and quality control.

It is a pleasure to thank Dr R. M. Haslam and other colleagues at the Royal Postgraduate Medical School for their efforts in this field and to state that our success is largely due to the expert assistance of Mr M. Longy and Mr B. Archer.

REFERENCES

- Blaivas, M. A., and Mencz, A. H. (1966). In *Automation in Analytical Chemistry*, p. 368. (Technicon Symposia, 1966, Vol. 1: papers from the Technicon Symposium held in New York City, October 17-19, 1966), Mediad, New York.
- Flynn, F. V., and Vernon, J. (1965). *J. clin. Path.*, **18**, 678.
- , Piper, K. A., and Roberts, P. K. (1966). *Ibid.*, **19**, 633.
- Gould, H. E. (1967). In *Automation in Analytical Chemistry*, p. 141. (Technicon Symposium, 1967: papers from the Technicon Symposium held in Brighton, November 13-15, 1967), Technicon International Division, Domont, France.
- Hoffmann, R. G., and Waid, M. E. (1965). *Amer. J. clin. Path.*, **43**, 134.
- Whitby, L. G., and Owen, J. A. (1965). *J. clin. Path.*, **18**, 668.
- Whitehead, T. P. (1966). *Hosp. Mgmt*, **29**, 708.
- Woodward, R. H., and Goldsmith, P. L. (1964). *Cumulative Sum Techniques* (ICI Monograph No. 3). Oliver and Boyd, Edinburgh and London.
- Wootton, I. D. P. (1965). In *Progress in Medical Computing*, p. 57. (Symposium organized by Elliott Medical Automation, London, 1965), Blackwell, Oxford.
- (1968). *Brit. med. Bull.*, **24**, 219.