

## 'On line' data handling in a routine haematology department

H. B. C. BURMESTER AND G. S. CROW

*From the Department of Pathology, Wythenshawe Hospital, Southmoor Road, Wythenshawe, Manchester 23, UK*

**SUMMARY** The use of a 'stand alone' mini computer, on line, data acquisition system is described utilising a microprocessor as an adjunct to the main central processor to increase data transmission rates. Visual display units (VDU) provide on line flexibility of data input.

The system acquires, on line, the Coulter S generated parameters together with the associated patient identification data and other requested tests. Erythrocyte sedimentation rate, platelet count, and B12 and folate results are fed in at will via the VDUs, using rapid entry subroutines.

Percentage differential white cell count results and the absolute leucocyte count are calculated by the microprocessor, on line input being facilitated by remote keyboards associated with the VDUs.

Total editing of all requests and results is available on all 'current' (unfiled) reports via the VDUs, and this facility includes the entry of film appearances using plus symbols and predetermined comments available against a code number. This editing function is also used to enter all other results that can be made available on the computer-printed report form.

The computer system also provides quality control data, immediate printout of Coulter S results outside given limits, cumulative filing for all routine and designated 'priority' results, and a daily alphabetical listing of all patients from whom specimens have been examined together with full blood count, ESR, and platelet results on specimens examined that day.

The constantly increasing workload in haematology led to the need for a change in the existing data handling and retrieval system, which consisted of a standard Coulter S card printout and manual filing (Nelson and Elder, 1972).

Quality control was being assessed on the result of a standard quality control blood examined at every 20th sample processed, and the results were reviewed as described in the ACP Broadsheet No. 75 (February 73) using the 'cusum' plot.

After a review of this system, it was felt that a radical change of the data handling procedures, utilising a departmental mini computer, was needed. The computer would also provide quality control information derived from all reported data including the 20th control specimen values. The system would be based on the 'on line' acquisition of Coulter S and WBC differential results; all other data, including patient identification data, would be committed to the computer via VDUs

with no transcription of patient identification data or results on to secondary media such as punch cards or tape.

Work study exercises had shown that 97% of historical reviews were carried out within 48 hours of the results being generated. Results would therefore be stored on a disc of such size as to permit approximately three months' work to be available for immediate historical review via VDUs. A 'priority result' facility was to be provided as well to hold designated results indefinitely.

The lack of experience in computer-driven, automated data handling techniques and limited financial resources determined that the installation of this form of 'on line' system was carried out in two separate stages.

The initial stage was the introduction of the new data handling procedure by means of a computer and associated hardware, which, however, provided only a limited element of computer automation by producing a strip report of Coulter S results, with NCR copy, as well as quality control information but leaving the full 'on line' all data

Received for publication 29 August 1978

acquisition and computer-generated report for the next stage.

A clerk was allocated to the haematology department to perform the clerical duties associated with the computer (for example, entry of patient identification data). This was found to be essential for the satisfactory running of the system.

**Review of initial stage**

Equipment: Automation Incorporated Alfa computer; 8K core memory  
Teletype  
Matrix printer

The Fortran program was written so as to provide a computer-generated six-digit laboratory number, the three-digit Coulter S number being ignored to avoid the too rapid repetition of the same number implicit in its use. The operation of the program was made as flexible as possible, patient identification, in the form of name and source of specimen together with the laboratory number, being entered on teletype tape before or after the acquisition by the computer of the Coulter S data. Reporting on either the teletype or printer could occur concurrent with, or subsequent to, the presentation of the specimen to the Coulter S using the laboratory number alone, if necessary, the results and quality control data being available at will from the computer.

Filing was by use of the back copy card of the request form as well as the NCR copy of the strip report of the Coulter S result.

The strip Coulter S report itself was attached to the request form, which provided the differential blood count and blood film report; this was then returned to the ward.

This limited procedure ran successfully for over 12 months and was followed by extension to the full 'on line' system.

**Equipment for full 'on line' system (Fig. 1)**

Computer: Capacity 32K words.  
Disc drive: 10 Megabytes dual disc drive.  
Microprocessor: Approximately 2K capacity.  
Visual display units: 6 VDUs with separate numerical keyboards.  
Teletype:  
Matrix printer:  
Coulter S interface:  
Program: The computer uses a Fortran 4 application program.

**Laboratory work flow for 'on line' system (see Figs 2 and 3)**

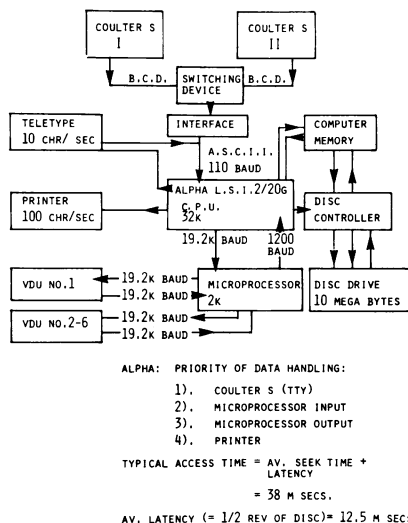


Fig. 1 Computer and peripherals (with data and control lines).

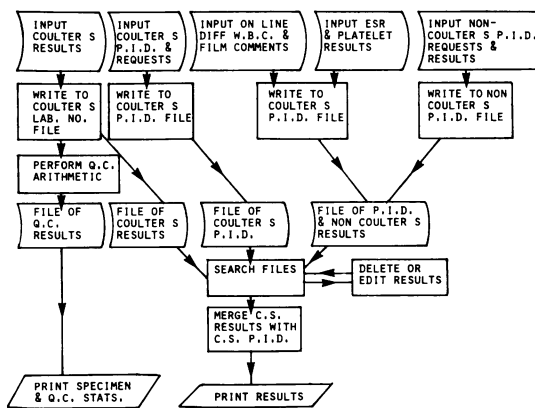


Fig. 2 Computer flow chart.

**1 PATIENT IDENTIFICATION DATA**

*Requests including Coulter S investigations*

All specimens from patients on whom a request for Coulter S parameters is received are assigned a unique six-digit laboratory number. This, together with the patient's name and forenames, sex, date of birth, hospital No., ward, initials of GP/consultant, and source of specimen, or a selection from these data, together with other requested tests, are entered into the computer via any VDU.

*Specimens without Coulter S investigations*

Specimens without an accompanying request for Coulter S parameters are not assigned a number

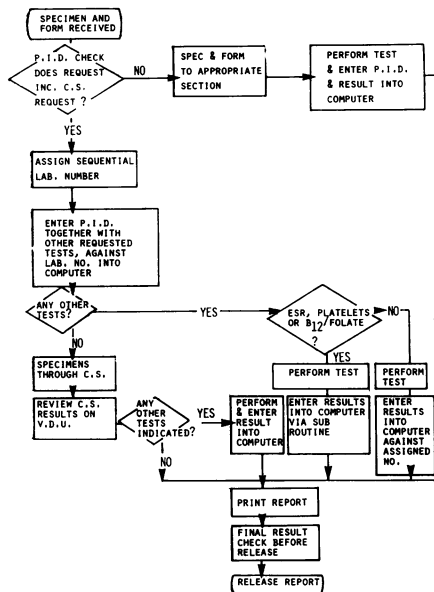


Fig. 3 Work flow diagram, 'Intelligent' Reporter (Medical & Scientific Computer Services Ltd, Lisburn, Northern Ireland), on line 32K system.

(thus keeping the continuity of the numbers of specimens being submitted to the Coulter S). Patient identification data (PID) are entered concurrent with/or before the entry of results.

### 2 SPECIMEN PRESENTATION TO COULTER S

Specimens are presented to the Coulter S in number sequence. The 20th quality control specimen is not numbered, the program having been written to expect quality control data at this interval, the sequence number 'running on' uninterrupted.

A facility is provided to examine out-of-sequence specimens and also to restart the number sequence from a given specimen number.

### 3 RAPID ENTRY SUBROUTINE

#### *Erythrocyte sedimentation rate and platelet count results*

Subroutines are available for entry of ESR and platelet results having a sequential laboratory number (that is, accompanying Coulter S results). These operate by presenting the facility to enter results sequentially from a given chosen sequential number.

#### *B12 and folate results*

These are entered singly or as a pair, depending on the tests performed, against full non-Coulter S PID when weekly batches of tests are completed.

### 4 REPORT EDITING

No report will print out until it has been reviewed on a VDU screen and all requests have been met. Outstanding requests are indicated on the VDU presentation of the current status of any series of investigations. All fields of the report may be edited with the exception of the laboratory sequential number. In addition, codes may be entered to indicate why a request has not been met, for example, CL = clotted.

### 5 DIFFERENTIAL WHITE BLOOD CELL COUNT

These are performed 'on line' by small remote keyboards associated with each VDU. The facility is provided to perform counts on 50, 100, and 200 cells. Absolute white cell type counts ( $\times 10^9/l$ ) are calculated by computer reference to total white cell count.

### 6 RETICULOCYTE COUNT

These are entered in the editing mode as a percentage count. The absolute count ( $\times 10^9/l$ ) is calculated by computer reference to the total red blood cell count.

### 7 FILM APPEARANCES AND OTHER INVESTIGATIONS

Commonly used red cell morphological characteristics (for example, anisocytosis, poikilocytosis) are quantified by the use of the plus symbol (+ to + + +).

A series of 64 predetermined comments on the characteristics of the film are available and are enterable as codes up to a maximum of six coded comments, for example, code 1 equals 'normal blood film'. Any other comments, as well as such investigations as RA factor, LE cells, prothrombin time, factor VIII, etc (see Fig. 4), are typed in using any one of the VDUs. All coded comments, but only one line of typed comment, are historically filed.

### 8 DATA SEARCH PROCEDURES

Data are held either as 'current' not yet filed or 'cumulative' reported and filed.

Current data are accessible either by laboratory number or by name with or without other PID.

Cumulative data are accessible by name with or without other PID.

Current searches present the full report format on the VDU screen. If necessary, the computer can be instructed to make a further search for a second, or alternative, report with the same PID if more than one specimen was collected in one day from the same patient.

Cumulative searches initially offer the option

WYTHENSHAWE HOSPITAL - HAEMATOLOGY DEPARTMENT

Name			Sex	D.O.B.	Hospital No.	Ward	G.P./Cons				
R.A.Factor	Paul Bunnell	HbS	L.E.Cells	Direct	Coombs	S.Folate	Vit. B <sub>12</sub>				
Screen		Screen		Test		µg/l	ng/L				
P.Time (secs)	A.P.T.T. (secs)	Thromb. Time (secs)	F.D.P. (µg/ml)	Bleeding Time (mins)	Factor XIII	Factor VIII					
Normal 11-14		35-45		10-12		<10		<7		60-150%	
Differential	NEUT.	LYMPH.	MONO.	EOSIN.	BASO.	ATYPICAL	BLASTS	METAMYELO	MYELOCYTES	NUC. R.B.C.	RETICS
W.B.C.	%	%	%	%	%	%	%	%	%	%	%
X10 <sup>9</sup> /L											
Comment Key	IN insufficient spec.		T.F. results to follow			N.A. results not available		C.L. Specimen Clotted			
Film appearance	Hypochromia		Anisocytosis	Poikilocytosis	Polychromasia		Macrocytosis	Microcytosis			

Investigations										
Date	W.B.C.	Hb.	R.B.C.	Hct.	M.C.V.	M.C.H.	M.C.H.C.	PLATELETS	E.S.R.	Lab No.
	x10 <sup>9</sup> /L	g/dl.	x10 <sup>12</sup> /L		fl.	pg	g/dl	x 10 <sup>9</sup> /L	mm/1 hr.	

Fig. 4 Computer-printed report form.

of selecting any line of the report meeting the given PID. This can be followed by a review of any full report selected. A return to any line or full report format can be made at any time during the search.

Data presented show the last report first and then in retrospective date order. A 'printout' of any report being cumulatively reviewed can be obtained.

A cumulative search involves a scan of both current and cumulative files.

The data presented after a name and PID search are given on all those filed records whose PID meet the given data, together with all other records whose PID do not exclude them from the search. Thus 'Smith' gives all the Smiths; 'Smith Fred' gives all the Smith Freds, together with all the Smiths without forenames. Indeed, 'S' will present data on all patients whose surname begins with S.

9 TELETYPE FUNCTION

This is used as an input/output device.

The accession numbers of out-of-sequence specimens can be input to allow immediate examination or re-examination of any specimens. Alternatively, the computer can be instructed to restart sequential

examination of specimens from any given accession number.

The teletype also serves as an input device to execute the various programs available to the users.

As an output device, the teletype prints the 20th quality control specimen results as they are generated, together with any abnormal results. Abnormal results are those values which fall outside the stipulated ranges for adult males, females, and children under the age of 10. In the event of sex and/or date of birth not being given, the program regards the patient as female.

Finally, the teletype prints certain messages related to the functioning of the system.

10 STATISTICAL DATA

Quality control

The computer provides mean and standard deviations on all Coulter S specimens. This information is available as a running population mean as well as the mean of the daily batch.

Mean and standard deviations are also available on a selected truncated subset of results, the trun-

cated ranges being the population mean  $\pm$  one standard deviation of the population mean.

The truncated range values are:

WBC	9.5 $\pm$ 4
Hb	13 $\pm$ 2
RBC	4.4 $\pm$ 0.7
Hct	0.39 $\pm$ 0.05
MCV	88 $\pm$ 6
MCH	29 $\pm$ 3
MCHC	32 $\pm$ 2

Further, the mean and standard deviation are available on the 20th quality control specimen result. The 20th quality control specimen results are excluded from other quality control calculations.

Accuracy is determined by assessment of the standard error of the daily mean against the population mean, and by the mean of the values for haemoglobin, red blood cell count, and mean cell volume on the 20th quality control specimen.

Precision is determined by the standard deviation of the 20th quality control specimen values for the same three parameters.

The performances of the two Coulter S machines are compared from time to time by estimating the standard error of the difference of the daily population means and performing a Student *t* test to establish the likelihood of each value belonging to the same population, using a null hypothesis.

#### Ministry statistics

The program provides the facility to count the source of each specimen and the total tests performed on a daily and cumulative basis.

#### 11 REPORTING

Reports can be generated at any time by typing in the appropriate code on any VDU. The computer will print reports on which all requests have been met and, in the case of a Coulter S specimen, have been reviewed before printing. Reports are routinely printed only once, but a further report is printed if that report is subsequently edited.

#### 12 FILING AND ARCHIVING

All current data are written to the fixed 5 mega byte disc of the dual disc drive unit. PID and Coulter S and non-Coulter S data are written to its individual files (Fig. 2). The final merger of Coulter S, PID, and results is performed immediately before the historical filing. Other historical filing is performed by the transfer of completed current results daily to the exchangeable 5 mega byte disc. Once historically filed, the reports are not available for editing. Incomplete reports are not historically filed but are

held over in an editable state; 700 incomplete results can be held at any one time on the fixed disc.

Results designated as of priority are 'flagged' and transferred to a priority result disc file weekly. The program provides for a priority code designating the haematological abnormality and subtype, for example, Anaemia : 1; Megaloblastic anaemia : 12; B12 deficiency : 121; Folate deficiency : 122. A search program is available to acquire data from the priority disc by use of this code.

The historical disc is archived weekly to maintain the integrity of the historical files in the event of a major 'crash' of the installed disc.

The total file is about 21 000 completed reports, representing approximately three months' workload. At the end of that time the oldest results are automatically culled by the program. Priority 'flagged' results are held for a further period of three months before culling. The results of the 20th quality control specimens are held on a separate file which is used to enable quality control calculations to be performed whenever required.

#### Discussion

When entering the automated data handling field the user's initial decision is whether to use a computer to implement the system of choice; implicit in making this decision therefore is the fact that one has, in truth, decided what the system of choice is going to be. It is possible, especially if one has had an unfortunate experience with computers, that an 'off the shelf package' is considered to be the system of choice. However, potential computer users may consider that they are probably in the best position to determine the system that would best suit their workload and wish to design a system of their own choice (Cavill *et al.*, 1974).

A critical decision in relation to the choice of system is whether an 'on line' or 'off line' submission of data to the central processing unit is chosen.

An 'off line' operation is cheaper in its use of hardware, less input terminals being necessary, but less cost effective in its use of time. Generation of fixed field punch cards or variable field tapes as an intermediate hard copy of the data, before batch processing, requires staff and staff time, which is not demanded by a 'real time on line' procedure; further, such procedures often require costly 'Date!' links to main frame computers.

To some extent, the 'on line' versus 'off line' decision is made for the user by the type of computer installation to be used (Cavill *et al.*, 1975). If an 'in the laboratory' mini computer is to be used only as an adjunct to a larger hospital or external (for example, regional board or teaching hospital) main



frame computer, then an intermediate hard copy of generated data may be the only means of passing such data to such a unit (Colvin *et al.*, 1977). However, if the laboratory mini computer is to be operated as a 'stand alone' system, then greater flexibility of choice may be possible (Alexander *et al.*, 1977).

Having decided on the system of choice, the writing of the program specification becomes the next operation. Within this the use and nature of PID is a major consideration, for on the method used will depend the ease of historical data searches. The choices appear to fall into two main categories. In a 'positive acceptance search', the program searches only for those patients who absolutely meet the given data, ignoring any other reports, even those of the same patient with which additional or less information is presented. This method has the advantage, therefore, of definitively identifying the required data. However, this is at the expense of the searcher having to know exactly which data were given for a particular request and then inputting a large amount of information to acquire the results. Further, a change of data, for example, moved from one location to another, or a change of spelling, can result in the data being lost within the system.

A 'negative acceptance search' creates the problem of a different patient's data being presented, if only limited data are offered for the search procedure. However, provided the name is given and the searcher is offered the opportunity of further selecting his information from data presented, a given report can be located by a less than complete input. An additional advantage of the 'negative acceptance' is the ability to include all patients in the computer files. This relates particularly to GP patients or emergencies who, lacking a hospital registration number, may otherwise be excluded from the system.

The speed of data transmission is crucial and is a prerequisite of any hardware purchase.

In the area of quality control, regimes are frequently based on the use of cumulative sum ('cusum') plots. However, difficulties were experienced in the use of these plots arising from minor variations in the day-to-day assigned mean values of the 20th quality control specimens. These variations, although within the manufacturer's performance limits of the machine, demanded continuous re-assessment of the assigned mean values.

The computer provided the facility to derive total population and truncated sub-set, and mean and standard deviation values, thus allowing the use of the standard error of mean calculation to determine daily accuracy.

In addition, the computer flags abnormal values and derives Ministry statistics.

The filing and archiving of data is an area in which seemingly little positive information, and much speculative assessment, exist. The initial selection of reports for historical filing obviously governs the number of reports held, and this in turn governs the size of the file required. In the system described, all reports are filed for a minimum of three to four months on a disc file large enough for the purpose. This procedure, when augmented with a priority category of reporting, which holds reports for six to eight months before culling and allows transfer of such results to a separate priority disc, offers a full archive for a period of time which, as the work study exercises have shown, is sufficient for the majority of historical searches. Priority results are assigned a code denoting the haematological abnormality present. The search program associated with such a code is seen as being of particular value for teaching and research activities.

### Conclusion

A stand alone 'on line' mini computer complete with data handling procedures is described, which by utilisation of an associated microprocessor, enables the department to be totally independent of all manual paper work. The 'on line' acquisition of data by VDUs provides a flexibility of input difficult to achieve by the use of either the conventional Coulter S card or other secondary medium.

A PID input procedure which accepts less than a totally comprehensive patient identification is coupled with a negative acceptance historical search pattern which, for historical reports, permits a latitude of data input while enabling individual records to be reasonably accurately and rapidly acquired.

In order to render operation of the system as simple as possible, a display format is produced on the VDUs which, as far as possible, is a replica of the final report, and the codes used for entering the various parameters are simple and easy to memorise. It is because of this that very little training is required to operate the system which can therefore be used by junior as well as more senior staff.

The system has been in operation for 16 months and, except for occasional hardware problems, has been completely satisfactory. The written program has required two minor alterations.

Medical staff in the hospital have commented on the excellent presentation of reports and the ease with which they can obtain information about previous results if a report has been lost. The

availability of computed absolute leucocyte and reticulocyte counts are of particular value.

Thus, in the department in which it is installed, this approach to computerisation has improved overall data handling as well as reduced wasted effort, and, by rendering reports more readily available and possibly more reliable, will, it is hoped, improve patient care.

We thank the haematology staff and all those who made it possible for us to have this system, in particular, Dr K. V. Lodge, Dr D. J. Brooke, and Mr E. Bett.

#### References

Alexander, M. K., Corbett, M. T., and Reed, R. W.

(1977). Maintenance and use of cumulative haematology files on a laboratory computer. *Journal of Clinical Pathology*, **30**, 356-360.

Cavill, I., Ricketts, C., and Jacobs, A. (1975). *Computers in Haematology*. Butterworths, London.

Cavill, I., Ricketts, C., Moulding, T., Jacobs, A., and Page, M. (1974). A system for data processing in haematology. *Journal of Clinical Pathology*, **27**, 330-333.

Colvin, B. T., Maynard, D., Henderson, A., Jenkins, G. C., and Fewell, R. G. (1977). Computer assisted data flow in a haematology laboratory (Abstract). *British Journal of Haematology*, **35**, 669-670.

Nelson, M. G., and Elder, E. (1972). Data processing in haematology. *Journal of Clinical Pathology*, **25**, 594-598.

Requests for reprints to: Dr H. B. C. Burmester, Department of Haematology, Wythenshawe Hospital, Southmoor Road, Manchester M23 9LT, UK.