

Clinical applications of an Automatic Data Handling System for sexually-transmitted diseases

GUNNEL ERIKSSON

Department of Dermatology, Södersjukhuset, Stockholm

AND

LARS PERSSON AND JAN WESSMAN

Research and Medical Departments, Astra Pharmaceuticals, Södertälje, Sweden

Interest in the sexually-transmitted diseases (STD) has increased during the last few years and the importance of studying the epidemiology of pathogens other than *N. gonorrhoeae*, such as *Trichomonas*, *Candida*, *Mycoplasma*, and *Chlamydia*, is becoming recognized.

In regard to gonococcal infection it is important to look for changes in the clinical picture, the antibiotic sensitivity of gonococcal strains *in vitro*, and the results of treatment in different areas. In this way such changes as an increasing frequency of asymptomatic infections or of pharyngeal gonorrhoea will be detected. Furthermore, treatment schedules can be altered as soon as a tendency towards decreased sensitivity *in vitro* for a particular antibiotic in a specific geographical area is observed or the frequency of treatment failures increases.

In a large VD department the amount of data grows very quickly and manual checking becomes time-consuming and impractical for the answering of some of the important questions mentioned above. It was therefore deemed necessary to construct an automatic data handling system.

We have now used such a system for more than 3 years, during which time it has been continuously developed. As the system has been found useful in its present form in providing answers to various questions, including those mentioned above, we thought it worth describing.

Main features of the system

General description

The system consists of one input program which also has the function of checking the material for errors and creating input for updating of the master file (Fig. 1). The master file is used as input for the data evaluation program of the system. This program has a main part and a variable number of data evaluation modules.

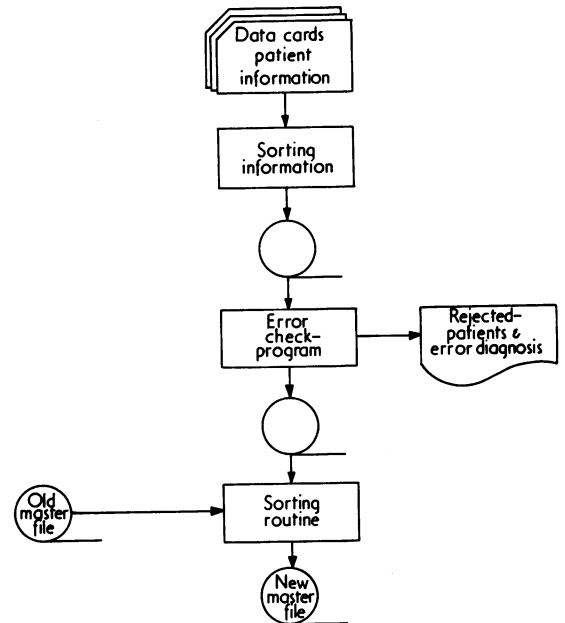


FIG. 1 System flow chart. Input part

The main program (Fig. 2) reads the parameter card, heading cards, and input master file, then selects the 'good' patients, *i.e.* the patients who satisfy the specified conditions, and further data concerning them for the actual evaluation modules. As soon as the main program reads EOF (end of file) on the input master file it starts the list module that prints the final output.

Technical details

The programs in this system are written in two different programming languages. PL/I optimizing compiler has been used for the error-checking program, the main program, and the list module of the evaluation program. All evaluations and statistical modules are written in FORTRAN IV.

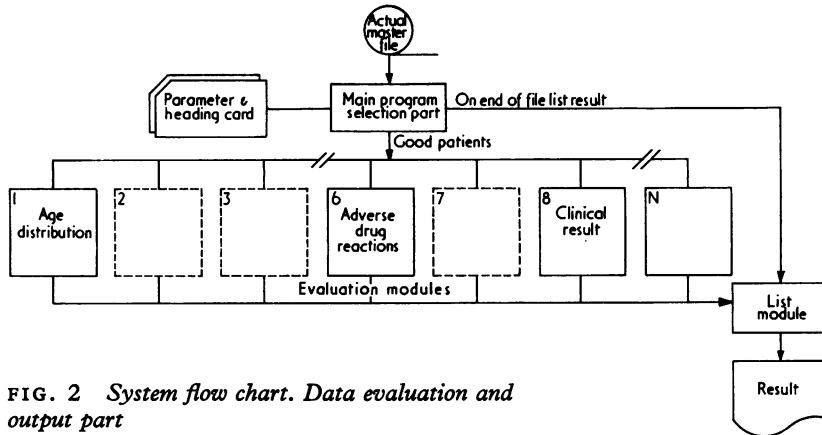


FIG. 2 System flow chart. Data evaluation and output part

The system is installed on an IBM 370/135 computer working under DOS/VS. The system uses 100 K words of real main storage and approximately 500 K words of virtual storage.

A typical running time is 4 to 7 minutes, depending on the size of the master file and the type of evaluation ordered.

Secrecy aspects

The technical staff, *i.e.* key-punch and computer operators, as well as the medical staff, is informed about the obligation to observe secrecy according to Swedish law.

All data are stored on tape after scrambling of the patients' birth dates. The tapes are stored in a separate archive to which only authorized personnel have access.

Form for recording patient data

In order to avoid extra editing, with an added risk of error, the patient forms have been designed in such a way that they can be used directly by the key-punch operator. Experience has shown that it is a quick and comparatively easy procedure to fill out these forms, providing that you have become familiar with the necessary code-lists.

Fig. 3 shows the standard patient record in which Part I deals with historical and treatment data, complications, treatment results, and side-effects; Part II deals with individual visits and mainly concerns data related to diagnosis.

Part I

Part I starts with the identification of the place of investigation *i.e.* clinic, hospital, and city followed by:

Identification of the study: transcript and study number (Nos 1-4).

Identification of the data card (5-6) identical with visit no. 11-12 in Part II.

Identification of the patient: running number (7-10), date of birth (11-16), sex (17), ethnic group (18).

Cause of visit: compulsory appearance, relapse, re-infection, probable re-infection, change of therapy due to complication, other reason (19).

Venereal disease history: previous venereal incident (20), preventive pills (21), serological reaction for syphilis (22), partner has symptoms of STD or diagnosed STD (23), open for own choice (24).

Treatment: drug (25-26), dose, mg. (27-30), doses/day (31), number of days treated (32-33).

Other medication: if yes, specified (34).

Kind and unit of sensitivity test (35-39), for each drug: MIC or IC_{50} in $\mu\text{g./ml.}$ or IU/ml., resistant-sensitive, zone diameter.

Complications of the disease (40-51).

Adverse reactions, including kind of reaction, no. of days after start of treatment when reaction appears, and association between investigated drug and recorded side-effect (52-72).

Result of treatment with respect to site of infection: cure, relapse, re-infection, probably re-infection, default, etc. (76-80).

Part II

Visit number (11-12) identical with nos 5-6 in Part I.

Gonorrhoea diagnosed marked only for the visit when the diagnostic test is obtained, *i.e.* when the positive smear or culture was taken (13).

Measure: therapy started, 1st follow-up, 2nd follow-up, etc. (14).

Date of visit (15-20).

Smear and culture sites: urethra, cervix/prostate, rectum, pharynx, other (21-32).

Site recorded: site from which specimen tested for sensitivity is taken (33).

Sensitivity *in vitro*: space for choice of five antibiotics (34-68).

Clinic:-----
Hospital:----- City:-----

Study no. Patient no. Date of birth Sex Ethnic group Cause of 1st visit

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

Prev. ven. incident Prev. pills Sph. serol. react Partner STD (Vacant) Drug Dose (mg) Doses/day Trmt. no. of days Other medication

Kind & unit of sensitivity test Complication of disease Adverse reaction

1 2 3

Result of treatment (site)

1	2	3	4	5
35	36	37	38	39

Day no.	Code		
40	41	42	43
44	45	46	47
48	49	50	51

Day no.	Association	Code				
52	53	54	55	56	57	58
59	60	61	62	63	64	65
66	67	68	69	70	71	72

1	2	3	4	5
76	77	78	79	80

Visit no.	CS diagnosed	Measure	Date of visit			Smear					Culture					Drug and sensitivity										Symptom			Complement-fixation titre			Menses			Sexual intercourse			Strain																															
11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80

FIG. 3 Standard patient record

Symptoms (69-71).
Complement-fixation titre (72-73).
Menses (78).
Sexual intercourse (79).
Strain (80): if two or more different strains are isolated at the same visit, successive numbers are used.

File creation

Standard 80-column key-punch cards have been used as input for the file creation program. These cards are punched directly from the patient record form described above (Fig. 3).

Before the cards pertaining to a certain patient are accepted as input to the master file, the input program checks several properties of the data. This error-check procedure includes the testing of conditions such as:

- All vital data being present.
- Only numerical data given in numerical fields.
- All cards pertaining to the specific patient being numbered correctly and having the same identification code.
- Specific control columns being empty.
- All patient visits being registered.
- Correctness of sex and treatment.

Any form not completely satisfying these criteria is rejected from the register for correction. When a patient is accepted by the program, the birth-date is scrambled to maintain secrecy and the data are stored in the master file.

Master file

This consists of patient records, each consisting of two main parts, one concerning mainly historical and treatment data and one concerning the individual visits the patient has paid to the clinic. These two parts of the patient file are in accordance with the patient form described above.

Computational procedures

Patient data selection

Before evaluation of a specific question, the patients to be included as well as specific data on these patients are selected according to pertinent criteria, defined on an *order-list*. The criteria are placed in a predefined order. The following is an example of such a list:

- Study No. 5.
- Ampicillin treatment.
- Males.
- Uncomplicated cases.
- Time limits: 1971-09-06 to 1972-09-05.
- Positive culture before start of treatment.
- MIC ≥ 0.1 µg. ampicillin per ml.
- Follow-up visit No. 1 within 10 days.
- Positive culture at follow-up visit No. 1.
- Denying intercourse since start of treatment.

Parameter card

This contains all the information necessary to guide the patient data selection and the choice of evaluation program. The selection criteria are given

in the predefined order mentioned above, and the pertinent data variables will be sorted through these during the execution of the corresponding evaluation program.

Heading card

The parameter card is always followed by a *heading card* when feeding the information necessary for the program execution. The heading card contains information mainly concerning the format of the data output and table headings.

Data evaluation

One of the main features of the system is its flexibility, which allows the inclusion of an unlimited number of further evaluation programs. To achieve full flexibility in the execution of these programs, they are all guided by the same type of parameter card, which is mentioned above.

TABLE I *Interval between start of antibiotic treatment of uncomplicated gonorrhoea and follow-up visits (Ampicillin 1 g. + 1 g. v. Doxycycline 0.3 g.)*

No. of Days	Visit number								
	2	3	4	5	6	7	8	9	
5	13	1	0	0	0	0	0	0	0
6	4	0	0	0	0	0	0	0	0
7	17	0	0	0	0	0	0	0	0
8	193	2	0	0	0	0	0	0	0
9	29	0	0	0	0	0	0	0	0
10	7	1	1	0	0	0	0	0	0
11	10	6	0	0	0	0	0	0	0
12	6	3	0	0	0	0	0	0	0
13	4	9	0	0	0	0	0	0	0
14	3	19	0	0	0	0	0	0	0
15	4	117	0	1	0	0	0	0	0
16	1	34	0	0	0	0	0	0	0
17	1	12	0	0	0	0	0	0	0
18	1	9	1	0	0	0	0	0	0
19	0	3	0	0	0	0	0	0	0
20	2	8	3	0	0	0	0	0	0
21	0	4	4	0	0	0	0	0	0
22	1	10	19	0	0	0	0	0	0
23	0	1	4	0	0	0	0	0	0
24	0	3	5	0	0	0	0	0	0
25	1	3	3	1	0	0	0	0	0
26	0	2	2	2	0	0	0	0	0
27	1	1	2	0	0	0	0	0	0
28	1	0	2	2	0	0	0	0	0
29	1	7	9	9	0	0	0	0	0
30	0	1	1	3	0	0	0	0	0
31	0	1	2	3	0	0	0	0	0
32	1	2	2	3	0	0	0	0	0
33	0	0	0	1	0	0	0	0	0
34	0	2	0	0	0	0	0	0	0
35	0	0	2	3	0	0	0	0	0
36	1	6	1	4	0	0	0	0	0
37	0	1	1	1	0	0	0	0	0
38	0	2	0	2	0	0	0	0	0
39	0	0	1	2	0	0	0	0	0
40	0	1	1	1	0	0	0	0	0
41	0	2	0	0	0	0	0	0	0
42	0	3	0	3	0	0	0	0	0
43	0	3	2	2	0	0	0	0	0
44	0	0	1	0	0	0	0	0	0
45	0	1	0	0	0	0	0	0	0
46	0	0	1	2	0	0	0	0	0
47	0	0	1	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0
50	5	14	7	9	3	3	1	0	0
Total	307	294	78	54	3	3	1	0	0

DESCRIPTIVE ANALYSIS MODULES

To obtain the main characteristics of a defined series of patients during a specific period of time, modules for a straightforward descriptive analysis have been constructed as follows:

- Age and sex distribution.
- Sensitivity of isolated bacteria.
- Complications of disease.
- Result of direct microscopy or bacteriological culture.
- Treatment results.
- Adverse drug reactions.

In Table I an analysis of the interval between start of treatment and follow-up visits is illustrated. The Table shows that most patients told to return after one week actually attended the clinic within 10 days.

LONGITUDINAL OR TREND ANALYSIS MODULES

In order to follow epidemiological changes in the micro-organisms, patients, and disease-characteristics, modules have been constructed to list pertinent data ordered according to suitable periods of time.

Trend analysis can then be applied to the data and significant changes over the time can be spotted. Examples of variables suitable for such analysis are:

Sensitivity of isolated micro-organisms to various antibiotics.

Failure rate with a standard antibiotic treatment.

Symptoms in patients with STD.

COMPARATIVE ANALYSIS WITHIN PATIENTS

In some instances it is of interest to obtain a detailed comparison between two variables. For this purpose a matrix plot program has been shown to be valuable. With this program the comparative sensitivity of gonococci isolated from a defined series of patient to two different antibiotics, such as benzylpenicillin and ampicillin, can be illustrated.

COMPARATIVE ANALYSIS BETWEEN PATIENTS

When comparing treatment forms, modules giving data in parallel patient groups have been used. The

computation illustrated in Table II can be taken as an example. These modules give tabulations of the antibiotic sensitivity of *N. gonorrhoeae* strains (in this case the zone diameter in the disc diffusion test) initially isolated in different patient groups. In Table II patients with positive cultures treated with ampicillin during three 4-month periods are presented. The module can be varied in different ways, e.g. presenting zone diameters or minimum inhibitory concentrations for the same strain for various antibiotics, or alternative treatment groups instead of different times.

Clinical applications

EPIDEMIOLOGICAL ASPECTS

Because of the growing interest in the epidemiology of STD we have tested the computer system on such problems. Questions which have been investigated are:

Symptoms in patients with gonococcal, candidal, or trichomonal infection or combinations of these infections.

Coexistence of such infections.

Sensitivity of *N. gonorrhoeae* to commonly used antibiotics, such as ampicillin and tetracycline, during consecutive periods of time.

TABLE II *Zone diameters for ampicillin found with paper disc method according to Ericsson, Högman, and Wickman (1954). Strains of N. gonorrhoeae from women treated with ampicillin during one year*

Tertial 1			Tertial 2			Tertial 3		
Zone	Frequency	Cumulative frequency	Zone	Frequency	Cumulative frequency	Zone	Frequency	Cumulative frequency
39	4	4	39	13	13	39	17	17
38	5	9	38	12	25	38	7	24
37	6	15	37	12	37	37	8	32
36	14	29	36	6	43	36	10	42
35	10	39	35	4	47	35	13	55
34	12	51	34	7	54	34	8	63
33	3	54	33	7	61	33	6	69
32	1	55	32	3	64	32	2	71
31	0	55	31	2	66	31	2	73
30	3	58	30	2	68	30	1	74
29	1	59	29	1	69	29	3	77
28	5	64	28	1	70	28	1	78
27	0	64	27	1	71	27	1	79
26	3	67	26	0	71	26	0	79
25	0	67	25	1	72	25	0	79
24	0	67	24	0	72	24	0	79
23	0	67	23	0	72	23	0	79
22	0	67	22	0	72	22	0	79
21	0	67	21	0	72	21	0	79
20	0	67	20	0	72	20	0	79
19	0	67	19	0	72	19	0	79
18	0	67	18	0	72	18	0	79
17	0	67	17	0	72	17	0	79
16	0	67	16	0	72	16	0	79
15	0	67	15	0	72	15	0	79
14	0	67	14	0	72	14	0	79
13	0	67	13	0	72	13	0	79
12	0	67	12	0	72	12	0	79
11	0	67	11	0	72	11	0	79
10	0	67	10	0	72	10	0	79

Number of patients without zones = 8

Investigated period 711101-721031

Total number of female patients = 262

EVALUATION OF DIAGNOSTIC PROCEDURES

For many reasons, including socio-medical and economic, it is important to re-evaluate the efficacy of the diagnostic procedures. Aspects of this problem which have been investigated are:

- Agreement between direct microscopy and culture.
- Comparison of findings in tests from different sites.
- Number of visits necessary for diagnosis.
- Number of visits necessary for follow-up.

EVALUATION OF TREATMENT REGIMENS

One of the problems in testing a new antibiotic drug is the often very insensitive test models, *i.e.* no good correlation can be found between the clinical and bacteriological results.

It has been shown that acute uncomplicated gonorrhoea can be used as a suitable test model (Eriksson, Magni, and Wessman, 1972). When this model is used in relation to a substantial number of patients a good dose/effect relationship can be established concerning clinical efficacy measured in terms of the disappearance of the gonococci. Also a good correlation between pharmacokinetic parameters and healing is seen, as well as between sensitivity of the actual bacteria and dose given. The computer system has been shown to simplify considerably otherwise very time-consuming evaluations of test data from such drug evaluations.

Summary

An automatic data handling system for the storage and evaluation of data from investigations on sexually-transmitted diseases has been developed. Among the clinical applications are epidemiological investigations, evaluation of diagnostic procedures, and evaluation of treatment regimens. One of the main features of the system is its flexibility, which allows the inclusion of an unlimited number of further evaluation programs. After a running-in period of about 3 years, during which time a number of changes were found to be necessary, the method has been found to satisfy medical and data processing requirements.

References

- ERICSSON, H., HÖGMAN, C., and WICKMAN, K. (1954) *Scand. J. clin. Lab. Invest.*, **6**, Suppl. 11, p. 23
- ERIKSSON, G., MAGNI, L., and WESSMAN, J. (1972) In 'Proceedings of the VIIth International Congress of Chemotherapy, Prague, 1971', ed. M. Hejzlar, M. Semonsky, and S. Mázak, p. 1401. Urban and Schwarzenberg, Munich

Applications cliniques d'un système automatique de traitement des informations dans les maladies transmises sexuellement

SOMMAIRE

Un système automatique de traitement des informations pour conserver et évaluer les données provenant des recherches sur les maladies sexuellement transmises, a été mis au point. Parmi les applications cliniques se trouvent les recherches épidémiologiques, l'évaluation des procédés diagnostiques et celle des schémas thérapeutiques. L'une des caractéristiques principales est sa souplesse, ce qui permet d'inclure ultérieurement un nombre illimité de programmes d'évaluation. Après une période de rodage d'environ trois ans, pendant laquelle il apparut nécessaire d'opérer un certain nombre de modifications, la méthode s'est montrée satisfaire les exigences médicales et celles du recueil progressif des informations.

Appendix

EXPLANATION OF SOME TECHNICAL TERMS

DOS/VS: DISK OPERATING SYSTEM/VIRTUAL STORAGE IBM Product.

EOF end of file: End of file routine: a routine which provides the special processing required when the last record of a file of data has been reached.

Executive program: A program which usually consists of a number of complex routines which reside wholly or partly in the main memory of a computer in order to monitor and supervise certain basic control functions.

Flowchart: The diagrammatic representation of a sequence of events, usually drawn with (conventional) symbols representing different types of event and their interconnexion.

Input: The process of transferring data, or program instructions into memory from some peripheral unit.

Key punch: A keyboard-operated machine used for punching data manually into punched cards or paper tape.

Key-punch operator: Person operating a key punch.

List module: Standardized module for printing of result.

Main program: The central part of a program which usually transfers control to other subroutines according to the nature of the data being processed.

Master file: A current, fully updated file to which new data can be added.

Matrix plot: A machine produced Cartesian coordinate system with conventional points substituted by frequencies.

Order-list: List of instructions.

Parameter card: A punched card on which the value of the parameters required by a program are punched and subsequently read.

Scrambling: System employed to secure secrecy of information.