PAPERS AND ORIGINALS

Computer-aided Diagnosis: Description of an Adaptable System, and Operational Experience with 2,034 Cases

JANE C. HORROCKS, A. P. McCANN, J. R. STANILAND, D. J. LEAPER, F. T. DE DOMBAL

British Medical Journal, 1972, **2**, 5-9

Summary

This paper describes a system of computer-aided diagnosis using an English Electric KDF9 computer linked to a terminal in a busy clinical department. Data from a series of patients were recorded, coded, and entered into the computer, which then performed a Bayesian analysis and displayed diagnostic probabilities in an adaptable format. Experience in this setting suggests that computer diagnosis may be a valuable aid to the clinician.

Introduction

We have elsewhere reported on some preliminary experiments in which we have studied the diagnostic process as practised by groups of clinicians and students (de Dombal *et al.*, 1969a, 1969b, 1971a, 1971b, 1972a) and also collated considerable amounts of clinical information into a relatively formalized "database" which might then be entered into the computing system (de Dombal *et al.*, 1971c; Staniland *et al.*, 1972).

In this and the following paper (p. 9) we describe a logical development of these earlier studies. The present paper outlines the essential details of the computer-assisted diagnostic system and briefly surveys the operational experience to date. In the following paper we review the results of a real-time prospective trial in which the computer's performance was compared with that of a group of clinicians.

System Description

The primary system shown in Fig. 1 involves the English Electric KDF9 computer situated in the electronic computing laboratory of the University of Leeds. This is about 800 metres from the university department of surgery, and is accessed via a Westrex ASR 33 teletype terminal located within the department of surgery itself. This system was utilized for the bulk of the diagnoses, and was in operation during the week from 9.30 a.m. until midnight. At other times a "back-up" system was used which employed a small desk-top computer, a Mathtronics 848 Biostatistician.

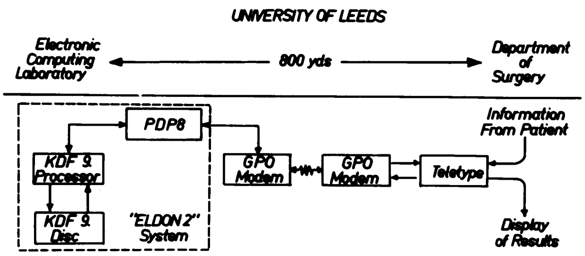


FIG. 1—Block diagram showing outline of primary system used.

Both primary and back-up computing systems utilized a variant of Bayes's theorem for analysing the patient data presented by each case entered into the system for diagnosis. The use of Bayesian probability theory in this form is not new; several previous attempts at such usage and a fuller assessment of the place of Bayesian statistical inference in medical diagnosis have been admirably set out by Lusted (1968). In the present instance the necessary programme to integrate Bayesian probability theory into clinical framework were specially written by two of us (A.P.M. and J.C.H.), as were the programmes for the display of information.

The ELDON 2 system which was used to provide access to the KDF9 computer has been fully described elsewhere (Wells *et al.*, 1971). It provides a means of storing and interactively amending both programmes and data held as a series of "files" on the KDF9 disc.

A computer "diagnosis" is made by placing an entry in a queue of jobs requesting the running of a compiled FORTRAN programme. This performs the Bayesian analysis on previously entered patient data and then prints out the terminal probabilities and information as requested. Depending on the number of other entries "in the queue" and the overall loading of the KDF9, servicing of this request can take anything from 30 seconds to 15 minutes.

Departments of Surgery and Computational Science and the Electronic Computing Laboratory, University of Leeds
JANE C. HORROCKS, Programmer, University Department of Surgery
A. P. McCANN, M.SC., PH.D., Lecturer in Computational Science
J. R. STANILAND, Research Assistant, Department of Surgery (Present address: Trinity Hall College, University of Cambridge)
D. J. LEAPER, M.B., CH.B., Research Assistant, University Department of Surgery
F. T. DE DOMBAL, M.D., F.R.C.S., Reader in Clinical Information Science

Mode of Operation

Fig. 2 shows the essential steps in any computer-aided diagnostic scheme of the type already outlined.

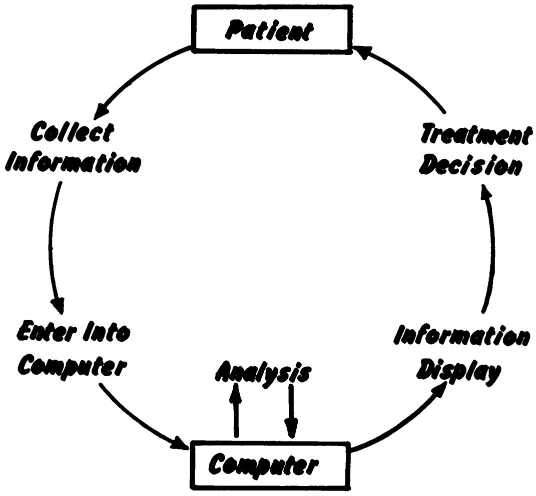
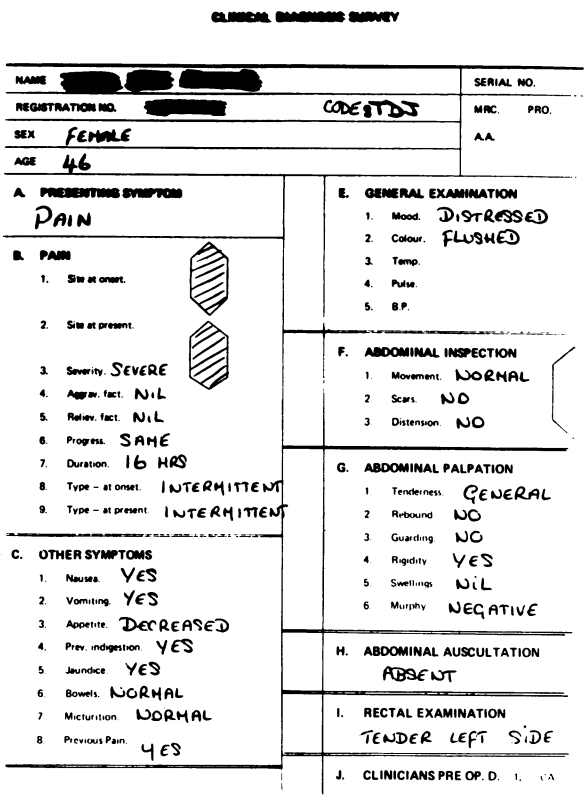


FIG. 2—Essential steps in any computer-aided diagnostic scheme.

COLLECTION OF INFORMATION FROM PATIENT

In advance of any computer-aided diagnosis we were obliged to create specialized forms in which the patient's case history could be to some extent "formalized," for later entry into the computer. An example of such a form is given in Fig. 3.

Even so, there is no guarantee that the data entered on even the most rigidly structured form will be "correct." We cannot claim to have solved this problem fully; all that one can say is that we undertook "observer variation" studies to investigate the subject which are reported elsewhere (Graham *et al.*, 1971; de Dombal *et al.*, 1972b). Furthermore, we allowed only persons to record data on the forms *after* they had undergone strict observer variation studies and in this way we have sought to minimize the problems of clinical data collection.



NAME [redacted]
SERIAL NO.
REGISTRATION NO. [redacted] CODE TDJ
MRC. PRO.
SEX FEMALE
A.A.
AGE 46

A. PRESENTING SYMPTOM
PAIN

B. PAIN
1. Site at onset.
2. Site at present.
3. Severity. SEVERE
4. Aggrav. fact. NIL
5. Reliev. fact. NIL
6. Progress. SAME
7. Duration. 16 HRS
8. Type – at onset. INTERMITTENT
9. Type – at present. INTERMITTENT

C. OTHER SYMPTOMS
1. Nausea. YES
2. Vomiting. YES
3. Appetite. DECREASED
4. Prev. indigestion. YES
5. Jaundice. YES
6. Bowels. NORMAL
7. Micturition. NORMAL
8. Previous Pain. YES

E. GENERAL EXAMINATION
1. Mood. DISTRESSED
2. Colour. FLUSHED
3. Temp.
4. Pulse.
5. B.P.

F. ABDOMINAL INSPECTION
1. Movement. NORMAL
2. Scars. NO
3. Distension. NO

G. ABDOMINAL PALPATION
1. Tenderness. GENERAL
2. Rebound NO
3. Guarding. NO
4. Rigidity YES
5. Swellings NIL
6. Murphy NEGATIVE

H. ABDOMINAL AUSCULTATION
ABSENT

I. RECTAL EXAMINATION
TENDER LEFT SIDE

J. CLINICIANS PRE OP. D. 1.

FIG. 3—Example of case history, showing form on to which patient data are copied for later entry into the computer.

ENTRY OF INFORMATION INTO THE COMPUTER

We allocated to each potential patient attribute a three-digit code number, and some examples of this are shown in Fig. 4. Instead of entering a full case history in English, it is now possible to inform the computer of the "facts" of a case merely by typing in a series of three-digit numbers, as is shown in Fig. 5. Incidentally Fig. 5 represents the case history of the patient in Fig. 3, and it requires little imagination to discern how the adoption of such a system greatly eases the problems of data entry into the computer.

CODING FOR EBIDMRC

SEX	
MALE	01
FEMALE	02
AGE	
0-9	03
10-19	04
20-29	05
30-39	06
40-49	07
50-59	08
60-69	09
70-79+	10

SITE OF PAIN	ONSET	PRESENT
RIGHT UPPER QUAD	11	22
LEFT UPPER QUAD	12	23
RIGHT LOWER QUAD	13	24
LEFT LOWER QUAD	14	25
UPPER HALF	15	26
LOWER HALF	16	27
RIGHT HALF	17	28
LEFT HALF	18	29
CENTRAL ABDOMEN	19	30
GENERAL ABDOMINAL	20	31

FIG. 4—Code for translating patient data (from form in Fig. 3) into three-digit numbers for entry into the computer.

```
CORRECT EBIDMRC
EBIDMRC 11FFF**OK
BL== **OK
INSERT==TDJ
002 007 020 031 061 038 044 046 049 052 055 063 065 067 069
071 073 078 082 085 087 089 093 096 100 102 112 115 117 118
121 123 125 127*
// **OK
!
NEW TEXT ON DISC
```

FIG. 5—Entry of case history into computer. These numbers each represent a clinical symptom or physical sign.

Moreover, clinical data can now be entered into the computer by any person familiar with the sort of simple code shown in Fig. 4. Thus the clinician need never enter into direct contact with the computer or even the computer terminal. Once a form setting out the patient's clinical attributes is filled out in English, the clinician need play no further part in the system until he reads the computer printout.

COMPUTER ANALYSIS

It has long been an ambition of those working in the field to feed in clinical data and allow the computer to select from its files the most appropriate diagnosis from the whole spectrum of recognized clinical ailments. Unfortunately this is not currently possible, and it is first necessary for us to assign a patient to a particular category—such as those indicated in Fig. 6. The diagnostic programme then collects the appropriate "database" about this class of diseases (which is stored as a file on the disc), performs a Bayesian analysis, and stores the resultant proba-

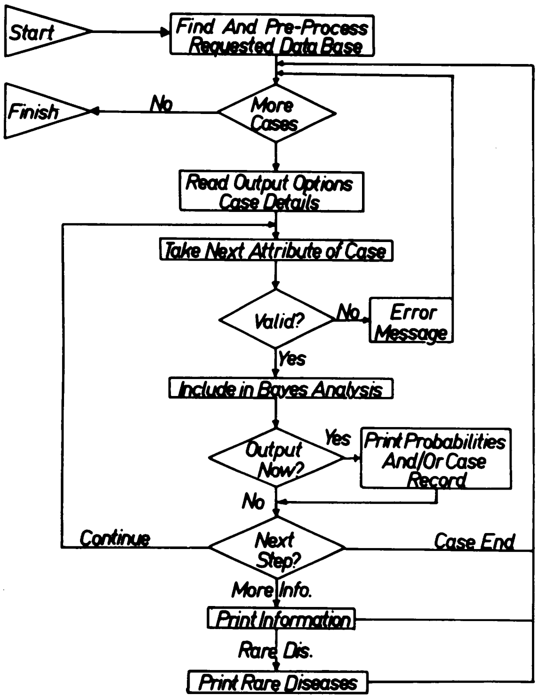


FIG. 6—Flow diagram of the steps in computer analysis and "diagnoses" of patient.

bilities. Note, however, that the computer is merely indicating that *if* the patient has one of the six or seven listed diseases in a "database" then the probabilities are as stated. Finally, the computer notes the type of display option requested (see below), complies with the request, and either moves on to the next case or finishes its analysis and displays the information via the teletype. Such an analysis uses about 7,700 words of core-store, and takes between 10 and 20 seconds to perform, depending on the number of cases entered and the precise display options ordered.

DISPLAY OF INFORMATION

We decided that at some stage a hard copy should be produced—this was necessary for documentation purposes. Secondly, we decided that this should be in colloquial English, rather than mathematical symbols, since it was likely to be read by busy practising clinicians, and we felt an obligation to present data to them in a readily comprehensible form. Next, we felt that the display system should be adaptable, since our earlier studies (de Dombal *et al.*, 1971a, 1971b) had indicated that there was considerable variation in the potential requirements of potential users of the system. Finally, we were acutely aware of the need for precautions to ensure the confidentiality of patient data entered into the system.

The type of display illustrated in Figs. 7, 8, and 9 fulfils these criteria. Firstly, the patient's name does not appear in the computer at all, merely an identifying code name and a series of code numbers. Secondly, the data are printed out as hard copy on the teletype. Thirdly, the computer starts by printing back the patient's case history (not only can the clinician readily understand this, but he can also check on any errors which may have crept into the coding of the history). Finally, the system can display information in a widely adaptable format.

Having printed out the case history and indicated the diagnostic probabilities on the information contained therein, the computer may then finish its analysis. However, two further display options are available. Firstly, the computer may compare its own diagnosis with that of the clinician. If, as in Fig. 8, these do not match, the programme attempts to select from filed information those most likely to resolve this discrepancy, and lists these attributes as suggestions for further checking. Secondly (Fig. 9), the computer may be called on to list rarer diseases, for which it cannot give precise mathematical probabilities, but which might just help the clinician in an obscure case.

```
CASE REF. TDJ
CURRENT SYMPTOMS
FEMALE
AGE............40 TO 49
SITE ONSET......GENERAL ABDOMEN
SITE PRESENT....GENERAL ABDO
SEVERITY .......PAIN NOW SEVERE
AGGRAV. FACT....NIL
RELIEV. FACT.....NIL
PROGRESS .......NO CHANGE
ONSET PAIN......12 - 24 HRS AGO
TYPE AT ONSET...INTERMITTENT
TYPE NOW.......INTERMITTENT
NAUSEA.........PRESENT
VOMITING.......PRESENT
APPETITE........DECREASED
INDIGESTION....PRESENT
JAUNDICE.......PRESENT
BOWELS.........NORMAL- NO CHANGE
PREV. PAIN.......YES-SIMILAR
PREV. SURG......NO PREV. ABD. OPRN
DRUGS..........NO TREATMENT
MOOD ..........DISTRESSED
COLOUR.........FLUSHED
ABD. MOVT.......NORMAL
ABD. SCAR.......ABSENT
DISTENSION .....ABSENT
TENDERNESS.....ALL OVER ABDOMEN
REBOUND .......ABSENT
GUARDING ......ABSENT
RIGIDITY .......PRESENT
ABD. MASS.......NOT FELT
MURPHYS SIGN...NEGATIVE
BOWEL SOUNDS...ABSENT
RECTAL EXAM. ...TENDER L SIDE
```

Fig. 7

```
POSSIBLE DIAGNOSES
 APPEND DIVERT PERFDU NONSAP CHOLEC SMBOBT PANCRE
PROBABILITIES ARE
   0·0    0·0    2·7    0·0    0·9    3·1   93·2
CLINICIANS DIAGNOSIS
   PRIMARY   -CHOLEC
   SECONDARY -SMBOBT
COMPUTERS DIAGNOSIS
   PRIMARY   -PANCRE  93·2
   SECONDARY -SMBOBT   3·1
NEITHER OF YOUR DIAGNOSES SEEM LIKELY. PROBABILITIES INDICATE
PANCRE AS PRIME POSSIBILITY
 ++ SUGGEST CHECKING THE FOLLOWING.......
AMYLASE
TENDERNESS....
SITE PRESENT
```

Fig. 8

```
++ SUGGEST CHECKING THE FOLLOWING......
AMYLASE
TENDERNESS..
SITE PRESENT
ADDITIONAL DISEASES WORTH CONSIDERING.....
PLEURISY
DIABETES
SYPHILIS
LEAD POISONING
SPINAL DISEASE
PERNICIOUS ANAEMIA
HERPES ZOSTER
O/A LUMBAR SPINE
```

Fig. 9

FIGS. 7, 8, and 9—Printout of case in Figs 3-5. For further comments on display options see text.

DECISIONS ABOUT THERAPY

From the computing system's point of view making decisions about therapy is the most straightforward aspect of the cycle. The computer makes no decisions at all about management. The computer's "responsibility" is thus severely restricted, being limited to (*a*) the provision of diagnostic probabilities for a subset of diseases, and (*b*) the recommendation of acquisition of additional information.

Operational Experience

This system came into being at the beginning of 1971. The Table sets out the overall operation experience since that date, from which it will be seen that just over 2,000 diagnoses have been attempted in conjunction with the computing system.

Summary of Overall Operational Experience: 2,034 Computer "Diagnoses"

Diagnostic Area	Retrospective Cases		Prospective Cases		All Cases	
	No. of Cases	% Accurate	No. of Cases	% Accurate	No. of Cases	% Accurate
Validation (500)*	—	—	—	—	100	96·0
Acute abdomen (600)	350	84·0	376	91·0	726	87·5
Lower G.I. tract (642)	300	79·7	82	87·8	382	81·4
Dyspepsia (175)	100	88·0	50	64·0	150	80·0
All areas	750	82·8	508	87·8	1,358	85·6
Estimates†	300	64·7	376	84·7	676	74·3
Total	1,050	77·6	884	86·5	2,034	82·3

* Figures in parentheses indicate numbers of patients in relevant "database" of information.
† Database created by asking group of six clinicians to supply probabilities about patients with acute abdominal pain.

The system has enabled apparently "real-time" diagnoses to be made. That is to say, we have in most cases been able to produce a set of probabilities from the computer in under 20 minutes, once data of the type shown in Fig. 3 are available to us. Indeed, on many occasions the time taken from data acquisition to probability printout has been much less, of the order of five minutes.

As regards reliability, this was one of our major potential problems, but in the event there was only one occasion when neither system was operative. (On that occasion data were obtained from the patients and stored for later, blind analysis.) Nevertheless, we would stress the need for careful integration of backup systems if "diagnostic computing" is to take its place in routine clinical practice.

Results

Our main experience (see Table) has been concentrated in the field of the *acute abdomen.* There is little doubt that the acute abdomen is a most suitable area for study in many ways; and in all some 376 diagnoses have been attempted. These are listed briefly in the Table, but the prospective controlled real-time trial which we have carried out is described in detail elsewhere (de Dombal *et al.*, 1972b).

We have also attempted to carry out studies in the more difficult field of *lower gastrointestinal disorders* such as ulcerative colitis and Crohn's disease. Here we encounter additional problems involving the nature of "diagnosis" itself. We have accepted histopathological diagnoses as being "correct," so that in a sense we are trying to predict what the pathologist will say rather than what is wrong with the patient. Nevertheless, it is of more than passing interest that our current prospective unselected series of 82 cases is being "diagnosed" by the computer with an accuracy rate of 88%.

Of considerable interest is the third area, which may loosely be described as *dyspepsia,* because this area was opened up entirely by three medical students in their spare time. Their experience too will be described in some detail at a later date; for now it might merely be worth remarking (*a*) that the "formalization" of clinical data necessary to enter such data into the computer is a highly useful experience for students, and (*b*) that it took less than three weeks, once the data were collated, for one of the students to create the necessary amending files, establish these on the KDF9, and begin to collect operational experience.

Finally, we attempted to create an information file based not on probabilities from known cases, but on *profiles*—that is to say, estimates of probabilities given to us by clinicians. This is an attractive concept in some senses, for if one enters a "clinician's thoughts" into the computer, then the clinician may well feel that he is exercising a considerable degree of control over the computing process. We originally undertook this task, since we were (needlessly) worried about the clinicians' reaction to the computing system, and since Edwards (1966) seemed to suggest that just such a system, in which probabilities are supplied to the system by means of human judgement, might have some advantages (Lusted, 1968). This was used to study the 376 cases of acute abdominal pain already mentioned.

Discussion

We originally opted for a simple flexible system, since we had in mind the grave dangers of creating a system which would work in the laboratory but not on the wards. In the event, the system proved perfectly feasible to operate. The diagnostic "turn round time" from collection of data from the patient to provision of probabilities by the computer was of the order of a few minutes only, perfectly acceptable in the clinical situation. The display formats proved easy to read, and sufficiently flexible to accommodate most clinical situations. The reliability of the interlocking systems has been such that on only one occasion were we forced to suspend activities on account of systems failure. Thus the system as envisaged and used is one perfectly feasible for use in a routine clinical environment, though we are acutely aware in saying this of the difficulties which face us in attempting to implement this (admittedly feasible) system as a routine clinical service.

The cost is remarkably difficult to quantify; but once created the system is relatively economical to run. The hardware involved (one teletype linked to the KDF9) costs around £500; the G.P.O. modem link can be hired for about £100-£150 a year, and the cost in central processing time (even had this been charged at full commercial rate) would work out at around £1 per "job." Remembering that up to 10 diagnoses can be performed in the same "job" (see Fig. 6) the cost per additional diagnosis can be as low as 10p. None the less, any attempt to cost our own system must take into account not only programming costs, which we are in no position to evaluate on a commercial basis, but also the cost of obtaining vast quantities of reliable and accurate clinical information. This we cannot cost accurately either, but as a guide it took some three man-years to prepare the "database" dealing with acute abdominal pain.

ROLE OF THE CLINICIAN

From the foregoing it will be apparent that the role of the clinician is undiminished in a system such as the one we have described. Indeed, in many ways it is enhanced. Thus the system is quite incapable of reliable operation unless the clinician first elicits reliable data from the patient—a curiously "old-fashioned" re-emphasis on the traditional values of accurate history-taking and careful clinical examination. Also the eventual "diagnosis," implying as the term does a decision about treatment for the patient, is left entirely in the hands of the clinician in this system.

What the computing system does is to help the clinician in an area where previous studies (de Dombal *et al.*, 1972a) have shown him to be relatively weak—namely, in the statistical analysis of large volumes of data. In such a case the clinician merely uses the computer to augment his own capabilities and judgement; and indeed there is ample precedent for this. To take one obvious example, the clinician often uses a stethoscope to augment his ability to hear sounds emanating from within a body cavity. Sometimes the clinical picture is clear-cut; in this instance the clinician merely uses his stethoscope to confirm his previous assessment of the patient. Sometimes the results which the clinician obtains from the use of the stethoscope are difficult to interpret or are at odds with what the clinician "feels" about the case—in such circumstances the clinician is at liberty to disregard the evidence from his "machine." But in a proportion of cases the evidence the clinician obtains will alter his impression of the case sufficiently to make him seek additional evidence and this in turn will lead him to the "correct" diagnosis.

This is precisely the type of computer usage which we have envisaged and tried to embody in the present system. No one speaks of a stethoscope making a diagnosis; and it seems to us meaningless to speak of the computer in terms which imply that this sort of machine system usurps the clinician's traditional role, even if, when the computer indicates its probabilities, we speak of the most likely complaint as being the "computer's diagnosis."

LEVEL OF COMPUTING EXPERTISE REQUIRED

Here the problem is much simpler. We have been impressed in earlier studies by the fact that clinicians are relatively reluctant and ineffectual users of any computing system (see de Dombal *et al.*, 1971a). The answer we have adopted is twofold. Firstly, we have instituted a three-tier system of computer training at undergraduate level, evidenced inter alia by the work referred to earlier. As regards the present we have instituted a computing system in which the knowledge and expertise required from the clinician is precisely nil. The clinician has merely to fill out a provided form, in English, and then subsequently to read a computer printout (again in English) as shown in Figs. 7, 8, and 9. The only "computer appreciation" necessary on the part of the clinician is an awareness that the computer is not infallible and that its "diagnosis" is merely an indication of probabilities on the data fed into it. Nothing else is needed, and certainly no programming or mathematical expertise is necessary for the clinician who elects to use the system.

It must be apparent from the foregoing that we owe a considerable debt of gratitude to many of our colleagues. It is a pleasure to thank Professor K. Smith, of the Department of Computational Science, and Professor M. Wells, of the Electronic Computing Laboratory, for their support and encouragement during the course of these studies. We thank also the members of their staffs who have given us much advice, comment, and helpful criticism, and without whom it would have been impossible to carry out 2,000 diagnoses in less than one year. Finally, three of us (J.C.H., A.P.M., and D.J.L.) were aided by a grant from the Medical Research Council, which we also acknowledge with gratitude.

Requests for reprints should be sent to Mr. F. T. de Dombal, University Department of Surgery, General Infirmary, Leeds LS1 3EX.

References

de Dombal, F. T., Hartley, J. R., and Sleeman, D. H. (1969a). *Lancet*, **1**, 145.
de Dombal, F. T., Hartley, J. R., and Sleeman, D. H. (1969b). *British Journal of Surgery*, **56**, 754.
de Dombal, F. T., Horrocks, Jane C., Staniland, J. R., and Gill, P. W. (1971a). *British Medical Journal*, **2**, 575.
de Dombal, F. T., Horrocks, Jane C., Staniland, J. R., and Guillou, P. J. (1971b). *British Medical Journal*, **2**, 578.
de Dombal, F. T., Horrocks, Jane C., Staniland, J. R., and Guillou, P. J. (1971c). *Proceedings of the Royal Society of Medicine*, **64**, 978.
de Dombal, F. T., Horrocks, Jane C., Staniland, J. R., and Guillou, P. J. (1972a). *Methods of Information in Medicine*, **11**, 82.
de Dombal, F. T., Leaper, D. J., and Gill, P. W. (1972b). Observer variation in recording the diagnostic process. In preparation.
de Dombal, F. T., Leaper, D. J., Staniland, J. R., McCann, A. P., and Horrocks, Jane C. (1972c). In press.
Edwards, W. (1966). *Transactions on Human Factors in Electronics* (*HFE*), **7**, 1.
Graham, N. G., de Dombal, F. T., and Goligher, J. C. (1971). *British Medical Journal*, **2**, 746.
Lusted, L. B. (1968). In *Introduction to Medical Decision-making*, p. 3. Springfield, Illinois, Thomas.
Staniland, J. R., Ditchburn, Janet, and de Dombal, F. T. (1972). *British Medical Journal*. In press.
Wells, M., Holdsworth, D., McCann, A. P. (1971). *Computer Journal*, **14**, 22.

Computer-aided Diagnosis of Acute Abdominal Pain

F. T. de DOMBAL, D. J. LEAPER, J. R. STANILAND, A. P. McCANN, JANE C. HORROCKS

British Medical Journal, 1972, 2, 9-13

Summary

This paper reports a controlled prospective unselected real-time comparison of human and computer-aided diagnosis in a series of 304 patients suffering from abdominal pain of acute onset.

The computing system's overall diagnostic accuracy (91·8%) was significantly higher than that of the most senior member of the clinical team to see each case (79·6%). It is suggested as a result of these studies that the provision of such a system to aid the clinician is both feasible in a real-time clinical setting, and likely to be of practical value, albeit in a small percentage of cases.

Department of Surgery and Computational Science and the Electronic Computing Laboratory, University of Leeds
F. T. de DOMBAL, M.D., F.R.C.S., Reader in Clinical Information Science
D. J. LEAPER, M.B., CH.B., Research Assistant, University Department of Surgery
J. R. STANILAND, Formerly Assistant in Surgical Research (Present address: Trinity Hall, University of Cambridge)
A. P. McCANN, M.SC., PH.D., Lecturer in Computational Science
JANE C. HORROCKS, Programmer, University Department of Surgery

Introduction

We have already described our general operational experience with an adaptable real-time computer-aided diagnostic system and discussed some of the problems inherent in its implementation (Horrocks *et al.*, 1972). But some important questions remain unanswered in the previous discussion. Chiefly these are two: do clinicians actually need such a system? and can it offer any measurable advantage in terms of diagnostic accuracy and reliability over more conventional methods? Such questions can be answered only at the bedside, in a routine clinical environment, by undertaking a controlled prospective trial in which the diagnostic performance of the unaided clinician is compared