Nov. 1967

42a

The problem of computer diagnosis

Computer diagnosis can already compete with the expert diagnostician in some restricted fields of medicine. The development of improved diagnostic programs combining comprehensive "knowledge" with the "experience" of many doctors is likely to give the computer an overwhelming advantage in the future. Dr. Reichertz describes his own experience in developing programs for the diagnosis of thyroid and other diseases.

Diagnosis is a vital and indispensable part of successful medical practice. In essence, it consists of comparison of the findings in a particular case with existing descriptions or the doctor's own memory of the features of different diseases. In differential diagnosis other possible conditions are considered in order of probability. Figure 1 shows in diagrammatic form the process of making a diagnosis, and indicates how the choice of special investigations is influenced by the history and clinical findings. The resulting diagnosis takes into account the experience of the diagnostician and sometimes the literature as well.

Diagnosis is thus dependent on 1. the quality of the data, 2. the knowledge and experience of the diagnostician, and 3. his ability to think out the diagnosis intelligently by recognizing the important features of a condition and making comparisons. Correct diagnosis can be expected only if all the relevant information is taken into account, and when the doctor considers all the possible diagnoses, and the appropriate part of his experience is not concealed beneath more recent impressions. It is immaterial that the process of association leading to diagnosis may have begun already during the history-taking. Actual diagnosis of an individual case begins only when the relevant information is available. Although rapid diagnoses made almost subconsciously are often falsely ascribed to intuition, this assumption overlooks the long training which precedes this ability.

Knowledge and experience

Figure 1 also shows that the findings in an individual case are compared in their turn with existing knowledge of the disease and with individual experience. With the almost exponential growth in scientific information it is no longer possible for the individual doctor to have a complete knowledge of even a subspecialty. Although such comprehensive knowledge is usually not needed in the everyday practice of medicine, it is none-theless interesting that tests have shown the average doctor's knowledge of symptoms and signs of even well-known conditions to be incomplete, while that of rarer diseases is almost always inadequate (Pirtkien, *Verb. dtsch. Ges. inn. Medizin,* 1966, 72).

This inadequate may be explained in part by the fact that the number of individual diseases and syndromes has increased from 5,000 to about 30,000 since the turn of the century, and will probably increase still further. However, the fact that the description of these conditions is not always sufficiently precise to allow for possible individual variations in symptoms and signs may also be partly responsible. It is therefore desirable in making a comprehensive diagnosis to have all the necessary information available.

The personal experience which has to be acquired by every diagnostician can be passed on to others only to a very limited extent. Like every scientific discipline, medicine strives so to organize knowledge that it becomes less dependent on individual experience—but this is easier to achieve in medical research than in diagnosis and practice. The possibility of combining individual and group experience of the frequency and variability of symptoms in different diseases may be advantageous to both practice and research.

Modern society often requires diagnostic studies of large cross-sections of the population, for instance in insurance medicine, public health surveys and the armed services. In large investigations of this kind, rapid evaluation of results considered from a single point of view, and free from subjective bias, has decisive advantages.

Fig. 1. Diagnosis-making as a flow-chart.
A picture of the patient's condition is obtained from the history, examination and results of routine and special investigations. This picture is then compared with known diseases and/or the diagnostician's own experience. The diagnosis so obtained is added to that experience and may also be published and thus added to the knowledge of the subject.

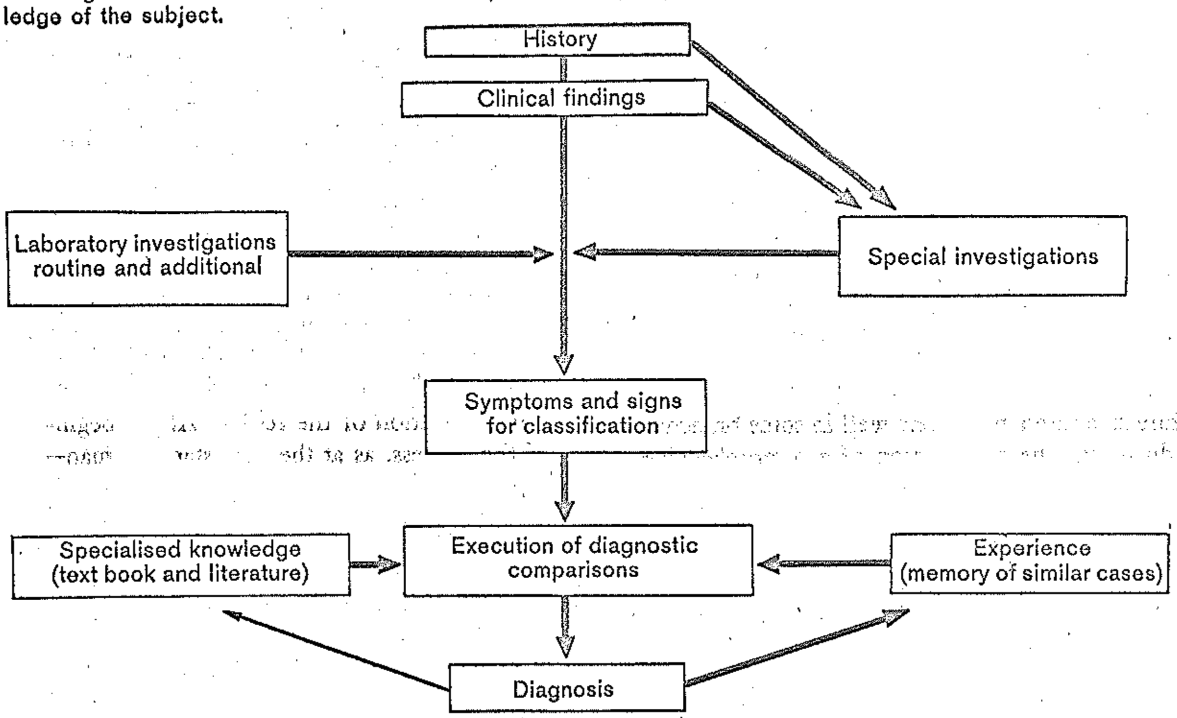
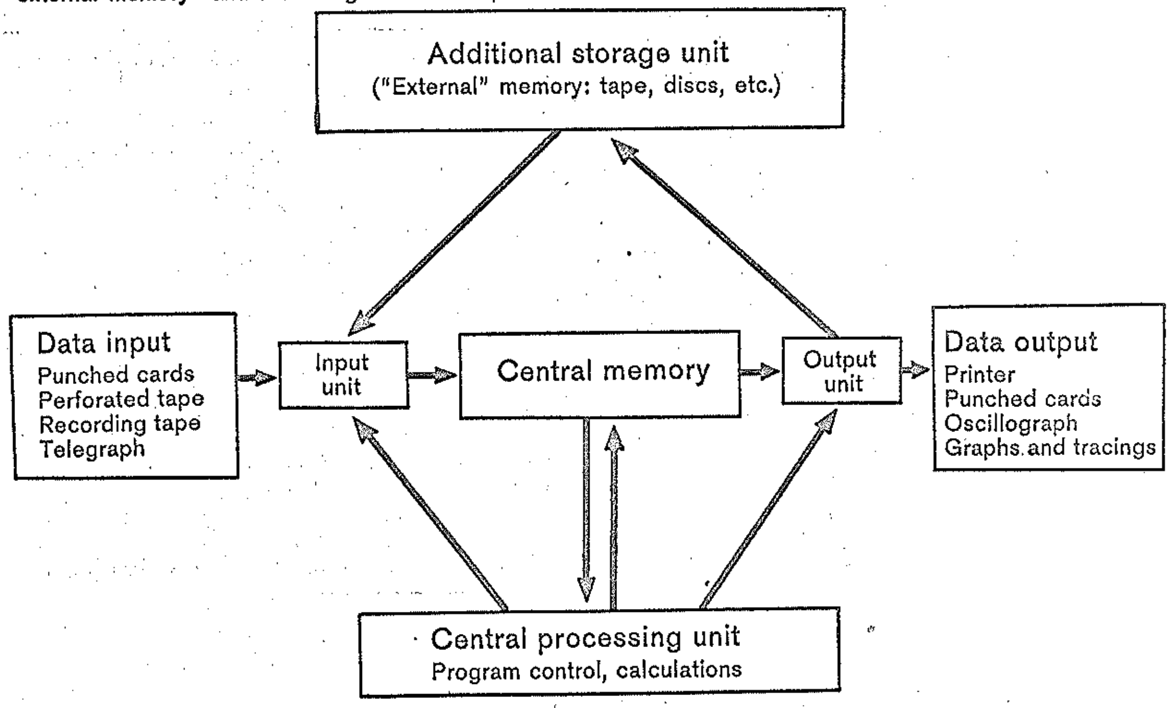


Fig. 2. Main working methods of a computer.
Information is fed in by various "data carriers" and is stored in the "central memory" together with instructions for handling the data (the program). The central arithmetic unit then carries out calculations and comparisons and the results are produced in accordance with the instructions given. Intermediate results can be stored in the "external memory" and fed in again when required.



The use of computers in diagnosis is justified by the need for detailed documentation of medical literature and disease entities, for accumulation in numerical form of the coordinated experience of many different workers and for the rapid handling of large quantities of data. Diagnosis can take its place alongside the existing computer applications in medicine—calculation and elaboration of statistical data, management or records, and hospital administration.

Figure 2 shows the main working methods of electronic data handling equipment. Programs (instructions for receiving data, handling it and giving results) as well as the data themselves can be fed in by various means such as punched cards and perforated tape. Both are stored in the computer's "central memory". Each instruction in turn is then applied to the stored data, and the results are again stored in the central memory, their issue also being controlled by the program. If intermediate results are produced in such quantity that the computer's memory cannot encompass them, then they can be temporarily stored on cards or tape in the additional storage unit, from which the information can be fed in again later in the calculation.

A computer cannot solve a problem on its own; if you feed in symptoms it cannot produce a diagnosis for instance! The method of solving a particular problem must be given in detail in the program, which must also include alternative instructions for dealing with specific situations which may arise in the course of a procedure. However, once the program is established the procedure can be repeated on different data with incredible speed and as often as required. Compared with a human physician, the computer thinks, so to speak, of everything that has been fed in and never "forgets" to take factors into account.

Programming

Programming for diagnosis is a prolonged and tedious but unusually fascinating task. It is concerned with fundamental diagnostic and logical problems and attempts to reduce the thought processes of the diagnostician to a formula and represent them mathematically.

The first stage is to sharpen up the definition of a disease by considering possible variations in symptoms and signs and the specificity of the latter in the particular case. Our own experience with programs for the diagnosis of thyroid diseases has shown that the program can be so designed that the definitions initially provided can be automatically improved as more and more data are fed in. In this way the accuracy of the program increases—the program "learns" by experience, and the improved definitions of disease which result are a useful by-product of computer diagnosis. Furthermore, the data on which diagnosis is to be based must be absolutely reliable and explicitly formulated ("hard"). The experienced diagnostician knows which findings may be unreliable or unimportant if he has made a diagnosis incompatible with his clinical impression of the patient; and although the computer initially regards all given information as having some weight, this does not present an insuperable difficulty. Firstly, it is possible to lay down "uncertainty factors" when feeding data and, secondly, the program can be so designed that it successively checks all data and puts the results at the disposal of the diagnostician. This is an important consideration in constructing our own program, but it involves a very large expenditure on programming and increases the computer time required.

Hard data (height, weight, age) involves no more than a slight probability of error. Unfortunately its diagnostic value is often relatively slight. As a basis for diagnosis, important facts such as the details of the history are liable to sources of error from both doctor and patient. In our own work on the program originated by Overall, Williams and Fitzgerald (Fitzgerald L. T. and Williams C. M., *Computer Diagnosis of Thyroid Disease.* Gainesville, 1965.) for diagnosis of thyroid diseases, we were astonished to find that different questioners got different answers from a patient to an apparently straightforward question, such as one about constipation or diarrhoea (Reichertz, Winkler, Kloss: *Verh. dts. Ges. inn. Medizin,* 1966, 72). The possibility of error can be reduced if the wording of questions is standardized and there are precise instructions for evaluating the answers.

The problem of handling details of the patient's history varies with the subject of the program and the range of possible diagnoses. For interpretation of electrocardiographs, a few unequivocal details like age, height and weight—or even none at all—are sufficient. Restricted diagnostic fields like thyroid or blood diseases require few details, mostly ascertainable in unequivocal form. The handling of larger subjects requires much more extensive information. Collen and his co-workers employ questionnaires covering the widest possible field in which the questions receiving positive answers are further subdivided in order to clarify the patient's answers. A very interesting development of this type of questioning may become possible when the doctor or his patient can communicate directly with the computer. The interrogation could then be more precisely aimed, each step depending on the previous answer so as to extract from a confusingly large number of available questions, only those expected to produce further useful information. Each question flashed on the screen would be answered by pressing the appropriate button of a keyboard—indicating "yes" or "no".

Fig. 3. The differential diagnosis of sore throat.
The individual findings lead to a branched program at the end of which either a precise diagnosis or diagnostic recommendations are produced.

Symptom: Sore throat; simplified block diagram of differential diagnosis

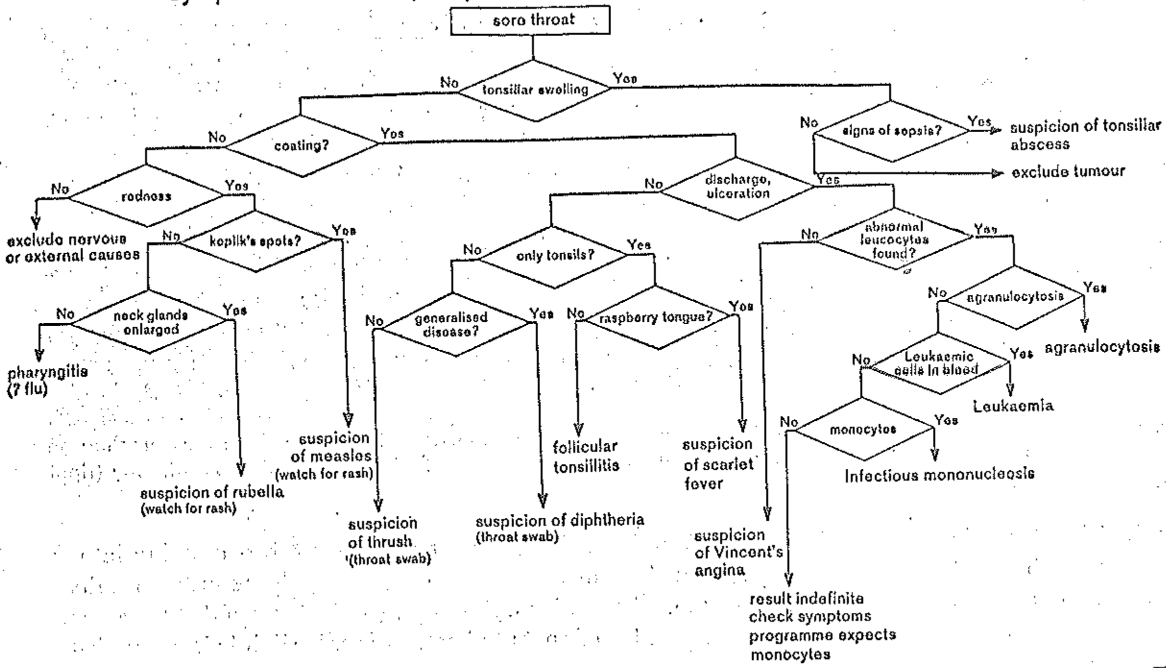


Fig. 4. Computer diagnosis of thyroid diseases.
This case report shows how a diagnostic classification of thyroid function was made on the basis of history and clinical findings alone, then on the results of laboratory tests alone, and finally using both together. The program included 17 clinical and 6 laboratory findings —not all of them essential to diagnosis. 10 and 3 respectively were employed in this case. After the functional diagnosis, a more general classification giving W.H.O. numbers is shown.
(Derived from program developed by Overall, Williams and Fitzgerald, cf. Reichertz, Winkler, Kloss: Dtsch. med. Wschr., 1965, 90, 2317).
(Calculations performed by IBM 7090/1410 computer at the Institute for Instrumental Mathematics, Bonn.)

Med. Polyclinic and nuclear med. dept. of Bonn University

Computer diagnosis of the thyroid (IV)
Patient 239

Type of data	Percentage Probability		Data employed
History and clinical findings	Hypothyroid	0	10
	Euthyroid	0	
	Hyperthyroid	100	
Laboratory tests	Hypothyroid	0	3
	Euthyroid	40	
	Hyperthyroid	60	
Combined	Hypothyroid	0	
	Euthyroid	0	
	Hyperthyroid	100	

Diagnosis	
No thyroid disease	0
Simple goitre (250)	0
Nodular goitre (251)	0
Hyperthyroidism (252)	100
Myxoedema and cretinism (253)	0
Other diseases (thyroiditis) (254)	0
Thyroid carcinoma (194)	0

Diagnosis

Problems so far discussed are only marginal to the actual making of a diagnosis by computer. Figure 3, which shows relevant possibilities and difficulties, is a simplified outline of diagnosis-making in the form of a flow-chart like those used in programming computers. The doctor often starts with a leading symptom or sign and tries to narrow the field in which this symptom or sign can occur by taking into account the presence or absence of other symptoms. Figure 3 is based on this principle and a diagnostic program could follow the same pattern. It would certainly classify many cases correctly, but some faulty decisions would be made since each finding leads to an unequivocal diagnosis. This can be a mistake since even typical features like the throat membrane in diphtheria can be absent from some cases, and others like Koplik's spots may go unrecognized (see Figure 3). Similar considerations can apply to most of the other features in the diagram.

In fact, only very few findings are decisive in diagnosis: evidence of suppression of radioactive iodine uptake by triiodothyronine excludes hyperthyroidism, and isolation of a particular organism can confirm a particular type of infection; but with the great majority of other symptoms and signs, the probability of their occurrence in several different diseases has to be considered. This is possible with the help of various mathematical models (see Reichertz. *Z. ärztl. Fortb.*, 1966, *55, 322*) which need not be described in detail here. Every such model so far produced has its own merits and limitations which have to be allowed for in practice, but they do have in common the ability to assign for each disease in the "repertoire" a figure indicating the possibility that the particular patient has this disease. Arranging these figures in order gives a diagnosis or differential diagnosis. Difficulties can be caused when several different conditions coexist, and it is then necessary to consider the possible combinations. This involves a considerable expansion of both the computer program and the capacity of the central memory. Increasing the diagnostic repertoire makes the machine decisions less definite and therefore less reliable. The number of diagnostic criteria must therefore be increased and in addition to their presence or absence their *specificity* must be taken into account.

Results

For these reasons, programs have so far been written for restricted fields of work like thyroid and blood diseases, congenital heart lesions, intestinal conditions and interpretation of electrocardiographs. Worthwhile results have been achieved. Warner used 33 possible diagnoses in his program for the recognition of congenital heart lesions and got results as good as those of an experienced cardiologist. Gustafsen reported similar results. In our hands, the program developed by Overall, Williams and Fitzgerald for diagnosis of thyroid diseases correctly classified 93 of 100 patients with hyperthyroidism. Agreement with expert diagnoses based on detailed laboratory investigations was reached in 95 % of cases when we introduced a "learning" function, which took into account each case already diagnosed when calculating probability data (see *Deutsche Medizinische Wochenschrift*, 1965, *90*, 2317). 95 of 105 patients were correctly classified on the basis of the history and clinical data, without B.M.R. or radioactive-iodine studies (*Verh. Dts. Ges. inn. Medizin*, 1966, *72*). Figure 4 is an example of a report from this investigation.

Collen and his co-workers successfully applied a computer program to the interpretation of prophylactic investigations performed at the rate of 20,000 per annum.

Pipberger, Caceres and Cady had good results with a program for interpreting electrocardiographs. One of their reports is reproduced in Figure 5. The diagnostic repertoire of this program, which is in course of development, is still rather restricted, but it gave correct answers in 34 out of 37 cases.

While computers have done well in some branches of diagnosis, the construction of a comprehensive diagnostic program still faces many of the difficulties already mentioned. These will be solved by improvements in computers and by the combined efforts of all the workers in this field. Our own aim is to assemble several independent sub-programs controlled by a main program, which would allot the task of detailed diagnosis to the appropriate sub-program according to the data provided. In every case the results of computer diagnosis will be a guide to differential diagnosis and further investigation—not a final and definitive verdict. This will remain the task of the doctor or his colleagues who alone see the patient as a psychosomatic whole, as a person, and carry the responsibility for his diagnosis and treatment.

Conclusion

When computer diagnosis has helped to reduce the chances of diagnostic error, it will have achieved an important part of his task. It has often been said that computer diagnosis may make medicine too materialistic and impersonal, but even if this were so an improvement in diagnosis would outweigh the objection. In any case it is not the computer which il ultimately responsible for computer diagnosis; its value depends on the quality of the program, on the investigation of the patient and on the interpretation of the results. At the beginning of the process, as at the end, stands a man—not a machine. What the computer has to offer is the exciting possibility of co-ordinating the experience of many experts and of making this available in practical form. It may be true that a little of the mystique of medicine will be replaced by realisme. But this would be a small price to pay for an improvement in the methods and scope of medical practice.

Dr. P. Reichertz
Lecturer in Medicine
Bonn University

Fig. 5. Interpretation of electrocardiograph.
This case report shows on line 1 the patient's name and date of birth, date of the investigation, registration details, the clinical diagnosis and the patient's stated age, followed by the serial number of the investigation. The diagnosis is considered under three headings and the final diagnosis is based on all three combined. The diagnosis of an old infarct is given no identifying letter, while "I" indicates an intermediate and "F" a recent occurrence. Further details of the tracing are given below the diagnosis.

48 Hoeltgen Lydia 1 7.0 11.0 28.0 23.0 5.0 66.0 200.0 59.4 200 right heart block 1 1 0 0 0 0 0 0 0 37.5 1

Diagnosis	Diagnostic (QRS) index	Percentage	Diagnostic (max. volt.) index	Percentage	Diagnostic (gradient) index	Percentage	Combined index	Percentage
Normal finding	29.02	7.61	25.95	6.99	0.01	0.01	18.27	4.88
	13.12	3.24	20.12	5.29	0.00	0.00	10.60	2.71
Anterior infarct	9.23	2.27	9.34	2.39	0.00	0.00	6.19	1.56
Anterior infarct F	20.81	5.26	21.90	5.80	0.00	0.00	14.23	3.70
Posterior infarct	53.58	17.36	47.71	15.05	3.58	6.30	34.87	12.89
Posterior infarct I	42.74	12.32	42.24	12.65	23.43	41.19	36.13	22.10
Posterior infarct F	6.76	1.65	7.37	1.88	0.00	0.00	4.70	1.18
Post. Lat. Inf.	58.67	20.44	52.39	17.44	0.00	0.00	36.94	12.60
Right heart bl.	66.20	26.42	67.47	28.86	29.86	52.50	54.51	36.01
Left heart bl.	13.84	3.43	14.16	3.66	0.00	0.00	9.33	2.37

Diagnosis (QRS)	Right heart block	**Diff. Diagnosis**	Post. Lat. Infarct Posterior infarct Posterior Infarct I
Diagnosis (max. volt.)	Right heart block	**Diff. Diagnosis**	Post. Lat. Infarct Posterior infarct Posterior Infarct I
Diagnosis (comb. grad.)	Right heart block	**Diff. Diagnosis**	Posterior infarct IPosterior infarct normal finding
Diagnosis (comb.)	Right heart block	**Diff. Diagnosis**	Post. Lat. Infarct Posterior infarct Posterior Infarct I

	QRS	max. volt.	Gradient	Combined
Combined	Yes	Yes	Yes	Yes

Other details of the tracing

Frequency	59.4/min.	????	−4.4 degr.	combined gradient 0.03 mVsec.
Angle of inclin.	−8.8 degr.	max. volt.	0.83 MV	angle at max. volt. −5.50 degr.