MEDICAL PRACTICE

Occasional Review

Computer interrogation of patients

R W LUCAS, W I CARD, R P KNILL-JONES, G WATKINSON, G P CREAN

British Medical Journal, 1976, **2**, 623-625

Summary

A system of routine interrogation of patients using a computer has been developed. It includes a visual display unit with a specially designed response keyboard, and the program has been designed to adapt to the individual patient. The system was evaluated objectively, using the criteria of accuracy in eliciting symptoms, acceptability to the patient, and cost. While doctors will always take the ultimate management decisions, it seems that machines can be programmed to undertake the routine interrogation of patients, elicit evidence accurately and acceptably, and calculate the probabilities of disease as effectively as doctors.

Introduction

Since the symptoms of a patient are often the earliest indication of disease, questioning, as a method of investigation, is likely to be necessary for a long time. If questioning by a computer program were as accurate as that by a doctor, if the method were as acceptable to the patient, and if the costs were comparable, a history might usefully be taken in this way.

Research Centre for Diagnostic Methodology, Southern General Hospital, Glasgow G51 4TF

R W LUCAS, BSC, PHD, research fellow and honorary lecturer, Department of Psychology, University of Glasgow

G P CREAN, PHD, FRCP, consultant physician and director

University Department of Medicine, Western Infirmary, Glasgow G11

W I CARD, MD, FRCP, senior research fellow

R P KNILL-JONES, MSC, MRCP, honorary senior registrar (hospital health services research unit)

G WATKINSON, MD, FRCP, consultant physician and gastroenterologist

The phrase "computer interrogation" is used to distinguish it from "history taking." In history taking the doctor elicits evidence from a patient by questioning and adapts the style and content of his questions to the particular patient since he receives numerous cues from the tone of voice, the facial expression, and gestures of the patient throughout the interview. These cues not only provide evidence about the reliability of the patient as a witness but also throw light on his attitudes and character. None of this accessory information is usually available to a computer, which can only question the patient and receive a reply.

Since Slack *et al* first reported its use in allergy[1] computer interrogation has been used in many other areas of medicine,[2-7] and some further development of particular systems[8,9] has been reported. Evaluation has been limited to discussion of the quality of the questionnaires administered[2,4,7] and to studies of patients' opinions.[2,4,5,7,10] The opinion studies have indicated a high level of acceptability. We describe here the results of the development and evaluation of a computer interrogation system.

Computer interrogation

In developing an "interface" for interaction between patient and computer one of our aims was to ensure that virtually all patients, including those with poor verbal ability and those with extreme anxiety, could interact successfully after a brief introduction. As the most acceptable form of interface will probably vary with age, sex, and intelligence, a second aim was the development of an interface that would adapt to the individual patient.

We found that significantly more patients (71%) preferred a visual display screen to a teletypewriter ($\chi^2=4{\cdot}97$; $P<0{\cdot}05$). Both types of terminal were about equally accurate in eliciting symptoms from the same patients, so we decided the visual display was preferable for computer interrogation.[11]

Extremely important in any form of self-administered medical questioning is the choice of works and phrases used in the questions. We therefore developed techniques for measuring the comprehension of medical terms,[12] and for choosing between several alternative phrasings of a question.[13] Only terms understood by 95% of the

population were used unqualified, and every attempt was made to use the simplest possible phrasing to improve both the accuracy and the acceptability of the method.

Also important in the presentation of questions is the speed at which words are typed on to the screen. So that the slowest readers feel unhurried, the slowest speed of most computer terminals, 10 characters per second, should be used. Faster readers feel frustrated by this, however, and their attention may wander. A decision algorithm was therefore developed for allocating patients to one of two speeds, 10 or 15 characters per second, according to age and verbal intelligence. In an experiment in which each patient's "ideal" speed was determined this algorithm produced only 5% misallocation to a faster than "ideal" speed.[13]

The three-button response keyboard developed at the National Physical Laboratory[3] has proved very successful in that virtually all patients seem to be able to use it. The three buttons allow "Yes," "No," and "Don't understand" responses, but many patients find the limitation to "Yes" or "No" too constraining. We therefore developed a keyboard that can be switched from this three-button format to a seven-button setting in which "Yes" and "No" answers are qualified by "Certainly," "Probably," and "Possibly." A decision algorithm was also developed for allocating patients to the three- or seven-button setting according to age and verbal intelligence.[13] From the response times to the introductory questions of the program, the verbal intelligence of the patient can be fairly well predicted ($r=0{\cdot}8$),[14] thus allowing the speed of presentation and the setting of the response keyboard to be adapted automatically to suit the individual patient.

Fig 1 shows a patient using the computer interrogation system.

FIG 1—A patient being questioned by the computer.

Evaluation

ACCURACY

It is convenient to think of interrogation as an information channel that connects events occurring within the body or mind of the patient to indicants (elements of evidence) written down on paper.[15] Accuracy can then be measured as the error rate of the channel—that is, of the doctor or machine interrogating the patient. The simplest channel is the binary channel where an event may or may not occur and may or may not be recorded. If an event occurs but is recorded as not occurring, the error is the false-negative or α-error. If the event does not occur, but is recorded as occurring, the error is the false-positive or β-error. Measurement of such error rates is difficult in medicine, since we may have no objective evidence whether the patient's experience of, for example, heartburn actually occurred. We can only observe agreements and disagreements between several independent interrogators, including the machine, and then, by some statistical model, estimate the error rates incurred by different questioners.

In an experiment of this kind by Card *et al*[15] 72 patients were each interrogated twice about 14 symptoms of dyspepsia, either by the computer and a doctor or by two doctors. The error rate ($\alpha+\beta$) of the computer was estimated to be 0·18 (18%) and the error rates ($\alpha+\beta$) of the doctors were 0·09 (9%), 0·09 (9%), and 0·12 (12%), the difference being significant ($P=0{\cdot}03$; one-tailed). In eliciting 76 indicants relating to alcohol abuse from 18 patients the error rates ($\alpha+\beta$) of two psychiatrists were estimated to be 0·14 and 0·16 and that of the computer 0·14.[16] In another experiment 36 patients were interrogated about the same 14 symptoms by two computer interrogation terminals—a teletypewriter and a visual display unit—using the same questionnaire. The two computers disagreed with each other on only 7% of all symptoms elicited, suggesting that data collected by these methods are highly reproducible.[17]

ACCEPTABILITY

While direct questioning of patients can provide useful information about patients' opinions, this information is subject to many sources of bias, and attitudes cannot be quantified in this way. The use of "attitude scales," constructed according to objective psychological scaling techniques, however, provides a method by which patients' attitudes toward computer interrogation can be studied quantitatively.

A modification of Edwards's "scale discrimination method"[18] was used to develop a scale for measuring patients' attitudes toward computer interrogation. The final scale consisted of 22 carefully evaluated attitude statements; the following are examples:

> Standards of treatment are sure to go up if computers are used a lot in hospitals.
> No two doctors would agree about what is wrong with you, so computers cannot do any harm.
> The computer will not be good enough to do any of a doctor's work.

To complete the scale patients were asked to indicate how far they agreed or disagreed with each of the statements. A second type of attitude scale, the semantic differential, was adapted for measuring patients' attitudes towards "medical interviews with a computer," "medical interviews with a doctor," and "the ideal medical interview." The semantic differential measures reaction to these concepts in terms of ratings on bipolar scales defined at their extremes by contrasting adjectives; examples are shown in fig 2. Both attitude scales were found to be reproducible and the good agreement between their measurement ($r>0{\cdot}60$) supports their validity.[19]

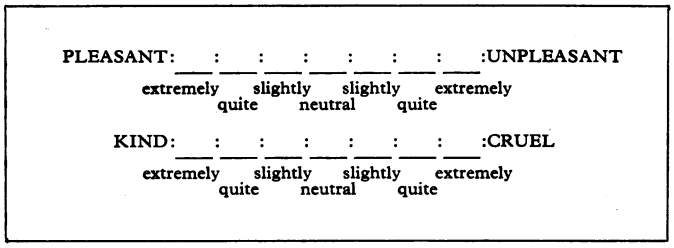

FIG 2—Some examples of semantic differential scale items used in measuring patients' attitudes towards computer interrogation.

After being interrogated by the computer about their dyspepsia 75 patients took a questionnaire containing the two attitude scales home to complete and return anonymously; 67 patients (89%) returned their questionnaires.

Altogether 55 (82%) patients had favourable attitudes towards computer interrogation. Men had more favourable attitudes than women ($P<0{\cdot}001$), patients aged 30 and under had more favourable attitudes than those aged over 30 ($P<0{\cdot}001$), and manual workers had more favourable attitudes than non-manual workers ($P<0{\cdot}05$). Forty-eight per cent of patients rated "medical interviews with a computer" better than "medical interviews with a doctor," and 49% rated "medical interviews with a computer" nearer "the ideal medical interview." The differences were again related to sex, age, and occupational category as above.[19]

Discussion

In assessing computer interrogation, it should be emphasised that it is routine clinical interrogation that is under study; we are not concerned with the complex interchange that occurs between patient and doctor in, for example, a discussion of a

personal problem, when a doctor has to continually respond to the many cues he receives from the patient. This method can probably be applied to many fields, but we emphasise the importance of selecting those questions that are most compelling in eliciting evidence and, also, of putting the questions in a form comprehensible to most of the appropriate population. This process may take many months if it is to be done effectively. For assessing the more limited form of interrogation described in this paper the criteria used are accuracy, acceptability, and cost. Our studies indicate that accuracy as measured by the error rates of the interviewers—the doctor and the computer—is closely comparable. Objective studies of patients' attitudes show that the method is acceptable to most patients, many of whom would probably prefer it as an initial way of collecting information. Until an ultimate system is established, however, it is difficult to make a clear comparison of cost. The present cost of interrogating in visual mode with a commercial time-sharing system is probably about the same as the cost of a consultant. In an ultimate system, however, a mini-computer with multiple terminals could be used, and this would be far cheaper. Large computers with on-line facilities are not required.

It is possible to foresee developments of computer interrogation. Usually questions are presented visually on a television screen, but experiments have been done with speech recorded on magnetic tape.[6] The use of speech should help those patients who are not fluent readers, and since regional differences in speech in Britain are considerable, questioning in the regional accent is likely to prove more acceptable. Possibly a combination of both modes, visual and spoken, might be best. In programmed learning[20] it has been shown that this combination is more effective than either mode alone. If a system of interrogation requires large numbers of questions with anything approaching random access to them, however, speech recording is unlikely to prove a suitable technique.

The chief difference in questioning between a doctor and a computer program is the doctor's ability to adapt; he is continually interpreting a stream of non-verbal information from the patient. The use of the response time of the patient to affect the setting of the keyboard and the speed of presentation is an attempt to introduce adaptive powers to the machine. The response time could be used to give other information. It has been shown that the longer a patient takes to answer a question presented by the computer, the more uncertain he is about his answer.[14] It should therefore be possible to estimate uncertainty in answers and thereby weight them, as indeed most clinicians do intuitively. Apart from the response time there are other characters of the patient—for example, pulse rate—that might be measured and used to provide information or influence the program.[8]

When a doctor questions a patient the patient's replies provide evidence for or against a set of diseases. There is no reason why the machine interrogation system should not be linked to a formal system of inference.[21–25] Given certain statistics about the set of diseases in the system, their changing probabilities could be calculated as interrogation proceeds and evidence accrues. The final set of probabilities could be made available to the doctor. A further development can be foreseen. At least some part of the physical examination of the patient might be carried out by paramedical staff and this further evidence used to recalculate the set of probabilities. It is then possible in theory, and might become possible in practice, to calculate what further tests—for example, radiological or laboratory investigations—could usefully be done. To use computer interrogation simply for collecting general medical data and presenting these to the doctor as lists of positive and negative symptoms fails entirely to exploit the power of this technique. The ultimate evaluation of any application of computer interrogation should be a formal trial against the traditional system of eliciting and using evidence, measured in terms of cost-effectiveness.

These new methods and the formalised medicine that is now possible force us to rethink our system of medical practice.[26] We shall require a partnership that combines most effectively the human qualities of the doctor, the qualities of the machine, and the discipline of decision theory. While doctors are highly skilled recognisers of clinical patterns and will always take the ultimate decision on the management of the patient, it looks as if machines can be programmed to undertake the routine interrogation of patients, to elicit evidence accurately and acceptably, and to calculate probabilities of disease as effectively as doctors.

This work was undertaken as part of a project funded by the Scottish Home and Health Department.

RPK-J was supported in the hospital health services research unit by grants from the Nuffield Provincial Hospitals Trust and the Scottish Home and Health Department.

We thank Dr C R Evans of the National Physical Laboratory for the loan of equipment.

References

1 Slack, W V, *et al*, *New England Journal of Medicine*, 1966, **274**, 194.
2 Mayne, J G, Weksel, W, and Sholtz, P N, *Mayo Clinic Proceedings*, 1968, **43**, 1.
3 Card, W I, *et al*, *Man-Computer Interaction*, p 141, Conference Publication No 68. London, Institution of Electrical Engineers, 1970.
4 Grossman, J H, *et al*, *Journal of the American Medical Association*, 1971, **215**, 1286.
5 Evans, C R, *et al*, *Computer Science*, report No 69. Teddington, National Physical Laboratory, 1973.
6 Marg, E, *et al*, *American Journal of Optometry and Archives of American Academy of Optometry*, 1972, **49**, 105.
7 Yarnall, S R, Samuelson, P, and Wakefield, J S, *Northwest Medicine*, 1972, **71**, 186.
8 Slack, W V, *Science*, 1971, **171**, 84.
9 Slack, W V, and Slack, C W, *New England Journal of Medicine*, 1972, **1**, 1304.
10 Slack, W V, and Van Cura, L J, *Computer and Biomedical Research*, 1968, **1**, 527.
11 Lucas, R W, *et al*, *The Applications of Electronics in Medicine*, p 161, Conference Proceedings No 34. London, Institution of Electronic and Radio Engineers, 1976.
12 Lucas, R W, *Social Science and Medicine*, in press.
13 Lucas, R W, PhD thesis, University of Glasgow, 1974.
14 Lucas, R W, *IRCS Journal of Medical Science*, 1974, **2**, 1176.
15 Card, W I, *et al*, *International Journal of Bio-Medical Computing*, 1974, **5**, 175.
16 Lucas, R W, *et al*, in preparation.
17 Lucas, R W, *et al*, *The Applications of Electronics in Medicine*, p 171, Conference Proceedings No 34. London, Institution of Electronic and Radio Engineers, 1976.
18 Edwards, A L, and Kilpatrick, F P, *Journal of Applied Psychology*, 1948, **32**, 374.
19 Lucas, R W, *International Journal of Man-Machine Studies*, in press.
20 Hsia, H J, *Journal of Communication*, 1968, **18**, 325.
21 Card, W I, *The Medical Annual*, p 9. Bristol, Wright, 1967.
22 Knill-Jones, R P, *et al*, *British Medical Journal*, 1973, **1**, 530.
23 Knill-Jones, R P, in *Ciba Foundation Symposium No 34 (new series)*, p 303. Amsterdam, Excerpta Medica, 1975.
24 Jennett, B, Teasdale, G, and Knill-Jones, R P, in *Ciba Foundation Symposium No 34 (new series)*, p 309. Amsterdam, Exceprta Medica, 1975.
25 Jennett, B, *et al*, *Lancet*, 1976, **1**, 1031.
26 Card, W I, *Proceedings of the Royal Society of London*, series B, 1973, **184**, 421.

What is the treatment for a four-months pregnant woman with osteochondritis juvenilis?

There are several sites for osteochondritis juvenalis, but presumably this refers to the disorder in the spine, also called Scheuermann's disease. It causes adolescent kyphosis of the thoracic spine and should give rise to no other problems in pregnancy than those of backache, which has to be treated on the usual lines—with rest on a firm bed, postural exercises, and possibly some form of spinal support. Because of the altered mechanics of the spine due to the disease and also pregnancy the backache may be felt not only at the site of the deformity but also in the lumbar spine. This needs similar treatment to that just outlined. Footwear should be checked to see that it does not cause any further distortion of the spine by undue tilting of the pelvis by inappropriate heels, and the length of the legs should be measured to see if any slight raise of one or other shoe might help to correct lumbar scoliosis.