PRACTICE OBSERVED

Practice Research

Effect of computer use in the consultation on the delivery of care

GARRY BROWNBRIDGE, ALAN EVANS, TOBY WALL

Abstract

The effects of the use of a computer on the delivery of care in consultations in general practice were examined. In this trial a computer system provided for the review and update of patients' medical histories, notes on doctor-patient contacts, and information on repeat prescribing. Thirty consultations in which the computer system was used and 30 consultations in which no computer was used were matched individually for the doctor consulted, the sex and age of the patient, and the presenting problem(s). Six independent general practitioners blind rated each consultation for the standard of care attained. A minor negative effect of computer use on the doctors' clinical performance was found. We suggest, however, that this small effect would disappear if a computer system was used routinely.

Introduction

Small, powerful, and reliable computer systems can now be bought at fairly low cost. Having a computer is a realistic proposition for the smallest business. What a computer does best is store, integrate, analyse, and retrieve information; such characteristics would be of benefit in general practice where accurate and fast processing of information might provide support for effective delivery of care. It is not surprising, therefore, that software packages are being developed for use in general practice. The interest of the profession is growing,[1,2] and schemes such as Micros for GPs, set up by the Department of Trade and Industry and organised by the Department of Health and Social Security, encourage general practitioners to explore the potential use of computers in general practice. This wide ranging scheme has sponsored the use of microprocessor systems in 140 practices in the United Kingdom and the effects are being evaluated. An interim report was published in 1984.[3]

As in the Micros for GPs scheme most computer systems have so far been used for administration. Typically, they provide facilities for patient registration, recall, screening, repeat prescriptions, and, more rarely, case notes. As the initial problems and upheavals in the practice routine are overcome, these new systems may open up opportunities—for example, in audit and preventive medicine—thus providing advantages over previous manual procedures.[4] The potential of such systems will be greater, however, when they reach into the consultation itself, enabling the doctor to retrieve from and add to the information system directly during the consultation—for example, by updating the repeat prescription file or adding to patient notes. This would integrate the administrative and clinical aspects of the doctor's work. Other potential uses include providing information (on prescribing, for example), prompts for opportunistic preventive measures, and providing treatment protocols to use for particular conditions.

Such benefits, however, will accrue only if systems are designed so that doctors can use them as an integral part of the consultation. More important, perhaps, is whether such systems carry unacceptable consequences for patients or doctors and for the provision of health care. Thus far research on the views of patients has shown that, although some dislike the idea of computers in the consultation,[4] the use of such systems is not resented.[5-8] Brownbridge *et al*, for example, concluded that "results indicate no major impact of consulting room use on patient views about the consultation. Their ratings of doctor attentiveness and rapport, information provision by the doctor, their own post consultation stress and satisfaction, and their intended compliance were not significantly

MRC/ESRC Social and Applied Psychology Unit, Department of Psychology, and Department of Community Medicine, University of Sheffield Medical School, The University, Sheffield S10 2TN

GARRY BROWNBRIDGE, BSC, research officer MRC/ESRC unit

ALAN EVANS, MB, MRCGP, general practitioner and associate in the department of community medicine

TOBY WALL, PHD, assistant director, MRC/ESRC unit

Correspondence to: Mr Brownbridge.

affected."[8] Doctors, however, have found that using such a system increases their stress, though only to a level equivalent to that experienced when a trainee is present or when faced with an unreasonably demanding patient.[7]

This finding raises perhaps the most important issue of all: whether or not the use of a computer system as an integral part of the consultation affects the standard of medical care provided. If such a system is stressful or demanding it may distract the doctor from his primary clinical task. More generally, interactive computer use is clearly the most intrusive way of deploying the new technology. This degree of interference may impair the doctor's delivery of care. As far as we know no investigation of this issue has been reported on. The present study attempts to redress the balance. First, however, it is necessary to consider how to measure the delivery of care in the general practice consultation.

There have been two main types of assessment of the standard of clinical care provided by doctors. The first, stemming from a research perspective, requires identifying a few specific illnesses or problem categories and creating a model or algorithm of the ideal clinical decision making procedure for dealing with them. The doctor is then assessed by comparing his or her performance against the prespecified criteria. The second approach, developed in the United States against the background of litigation minded consumers and defensive medicine, is often referred to as "quality control in medical care." This entails retrospective analysis of medical records only. Both types of medical care evaluation have been reviewed.[10-13]

For our purposes, however, neither of these approaches is suitable. In the first case it is impracticable to develop a sufficient number of algorithms to cover the range of problems encountered by the general practitioner. Alternatively, limiting the assessment to a few conditions would result in examining the effects of computer use on a small and atypical sample of the doctor's work. The second approach, based on written records only, similarly limits the inquiry to only a small part of the clinical process. Although in the USA, in case of litigation, the medical records must be a complete and comprehensive record of the delivery of clinical care, this is not so in Britain. The general practitioner records only information that he or she considers to be relevant to clinical decision making and management or for future reference.

What is required is a method that assesses the total encounter between the doctor and patient and covers the breadth of the presenting problems of a typical surgery. Pendleton *et al* provide a framework within which general practitioners and their trainees may consider the effectiveness of their consultations.[14] They propose a method for evaluating a consultation by applying a consultation rating scale. This general approach was put into practice for this study.

Method

BACKGROUND

The study was based on the use of a computer in the consulting room by five doctors in a general practice with 20 000 patients on its list, working from two sites in a predominantly working class area of Sheffield. The patient populations served by the two sites were equivalent and largely homogeneous with respect to ethnicity and social class. The computer system was developed jointly by IBM (UK) Scientific Centre and Sheffield University Medical School to provide a comprehensive range of administrative and consulting room facilities. It has been fully described elsewhere.[15] The system could be used during consultations through terminals on the doctors' desks to call up the patient's recorded medical history, to update the notes, and to change the medication file for repeat prescriptions. This facility was available at one site of the practice but not at the other. One site thus provided a control. All of the doctors practised from both sites.

As part of a larger study 89 video recordings were made of all consultations in a selection of surgeries at both sites. These covered roughly 180 consultations during which the computer system was used and over 550 consultations in which the computer was not used (before it was installed or at the control site). Early analyses revealed no discernible effects of video recording on patients' reports of the consultation[16] and no difference between consultations over time. These video recordings provide the data bank on which the study was based.

RESEARCH DESIGN

Thirty consultations in which doctors used the computer were individually matched with consultations from the sample in which a computer was not used. Matching was exact with respect to the doctor seen, sex of the patient, the age of the patient (within 15%), whether it was a first visit or follow up visit, and the type of presenting symptoms. Matching for the last variable was carried out with the supervision of a doctor by a trained nurse using the classification system developed by Raynes (table I).[17] Because these are broad categories further medical judgment was also used so that

TABLE I—*Raynes's categories of patients' presenting symptoms*[17]

(1) Physical symptoms
(2) Psychological symptoms (mood, behaviour, depression, anxiety, marital problems, delusions, phobias)
(3) Social problems (housing, leisure, work or family problems, excluding marital problems)
(4) Patient feels unwell but is unable to identify specific symptoms
(5) Problems relating to pregnancy
(6) Improvements in condition being treated
(7) No change in, or worsening of, condition being treated
(8) Physical symptoms with psychological symptoms superimposed
(9) No symptoms

problems of obviously differing severity were not matched. As computer use was known to affect the duration of the consultation no attempt was made to match on this factor.[9] The sample was also selected to represent equal numbers of consultations by each of the participating doctors, and, for each doctor, to represent as broad a range as possible of different symptoms. Sixteen pairs of matched consultations were follow up visits, with reasons for attendance including, for example, nappy rash, bed wetting, viral infections, ulcers, haemorrhoids, diabetes, depression, and postoperative care. Fourteen consultation pairs were first presentations of symptoms, such as coughs and colds, aches and pains, urinary problems, rashes, lumps, and lesions. Five consultation pairs were for multiple problems, with equivalent combinations of symptom types in each consultation. The research design thus called for a comparison between computer use and non-use conditions in a carefully matched sample of typical general practice consultations. The standard of care attained by the doctor was rated blind, using the procedure described below.

MEASURING DELIVERY OF CARE

To measure the standard of the delivery of care an independent panel of six principals in general practice, all of whom were also concerned in a well established vocational training scheme (four of them were trainers), designed with us a delivery of care rating scale in a series of four discussions, each lasting two to three hours. The panel chose to design their own instrument rather than use the one described by Pendleton *et al*, as the emphasis of that instrument did not accord completely with the panel's views of what was good clinical performance during a consultation. Their brief during the discussions was: to identify the relevant dimensions on which consultations should be assessed; to decide on appropriate items and response format for the rating scales; and to agree a final form when it had been applied to a pilot sample of video recorded consultations. One of the first decisions made was to draw a distinction between the appropriateness and adequacy of a doctor's clinical behaviour. These two aspects were then related to seven areas—namely, the doctor's performance in: (*i*) identifying the relevant complaint(s); (*ii*) identifying relevant background factors; (*iii*) conducting a physical examination; (*iv*) interpreting the findings from the verbal and physical examinations; (*v*) reaching a decision on the outcome of the consultation (including prescriptions, plans for further investigation, recall, and referral); (*vi*) communicating with the patient; and (*vii*) making a record of the encounter. A final item required rating the doctor's efficiency in the use of time. To each of the above variables was attached a seven point scale, running from inappropriate/inadequate (scored −3) to appropriate/adequate (scored +3) (mid point 0). The resulting 15 item instrument is available on request.

Having thus designed and piloted the instrument, the panel then applied it to the matched consultations. Each rater was required to assess 15 pairs of consultations. Thus each was provided with the notes available to the doctor at the time of the consultation, a complete verbatim transcription of the consultation, a full description of the physical examination conducted, and the notes made as a result of the consultation. The transcript was of all the utterances of doctor and patient. If the doctor mentioned either the computer or the notes this was changed to read "records." During silences the doctor's activities and their duration were described. Use of computer or paper records were both referred to as "use of records." This procedure was necessary to preserve patient anonymity and to ensure that the rating was

carried out blind. Consultations were presented to the raters in random order and in such a way that each case was rated independently by three separate raters.

In rating procedures such as that used in this study it is desirable not only to have a demonstrably expert panel but also to show that individual experts agree in their assessments. If the delivery of care is a measurable aspect of a doctor's behaviour then it is necessary to show that it can be consistently quantified. Thus the panel's ratings were examined for interrater reliability across consultations.

Over all consultations, and on each item, the distributions of each rater's scores were highly skewed (towards the positive end of the scale), making the use of parametric correlation statistics inappropriate for the assessment of interrater agreement. Non-parametric tests of correlation were also inappropriate as too large a proportion of ratings was tied. For all scale items individual raters often assigned the same score to several of the consultations they rated. Therefore, percentage agreements were used. The criterion set for agreement between any two raters was that there should be no more than a single scale point difference between them when rating the same consultation on the same item. Using this criterion we found that across all pairs of raters over the 15 items agreement was typically 60% to 70%; mean per cent agreement scores ranged from 45·6% to 80% (see table II). With the seven point scale used and agreement defined as plus or minus 1, agreement owing to chance alone would be expected in 19 (38·8%) of 49 instances.

TABLE II—*Effects of computer use on the delivery of care*

Delivery of care item	Interrater agreement (mean %)	No computer use (mean rating)	Computer use (mean rating)
Identified complaints:			
Appropriately	75·5	1·8	1·6
Adequately	63·9	1·4	1·3
Identified background factors:			
Appropriately	68·9	1·2	1·3
Adequately	62·7	0·9	1·0
Conducted physical examination:			
Appropriately	79·5	1·4	1·2
Adequately	70·6	1·0	1·0
Interpreted evidence:			
Appropriately	78·9	1·6	1·5
Adequately	64·5	1·5	1·2
Medical decision reached:			
Appropriate	80·0	1·9	1·7
Adequate	68·9	1·5	1·5
Communication with patient:			
Appropriate	65·5	1·6	1·3
Adequate	45·6	0·9	0·7
Records made:			
Appropriate	68·3	1·5	1·3
Adequate	57·2	1·0	0·6
Use of time:			
Efficient	56·6	1·4	1·1

Using this level of agreement that would occur by chance as the base line, the χ^2 statistic shows that agreement above 60% is significant beyond the $p<0·001$ level (and 56% agreement is significant beyond the $p<0·002$ level).[18] Clearly the raters in this study showed considerable convergence in their assessments of the standard of the delivery of care across consultations. Interestingly, the worst areas of agreement concerned the adequacy of communications with patients and of records made. Perhaps doctors are not the best judges of the adequacy of communication, whereas it seems reasonable to assume that the latter outcome—records made—reflects controversy in the profession concerning how selective or all-inclusive doctors should be in making patient notes. The overall point, however, is that this procedure produced an appropriate, valid, and reliable measure of important aspects of the delivery of care for use for research.

Results

The objective of this study was to determine whether the use of the computer during the consultation had any discernible effects on the delivery of care. The relevant results are shown in table II, where expert ratings of the standard of care attained in 30 conventional consultations are compared with those for 30 matched consultations during which the computer was used by the doctor. Correlated sample *t* tests show that item by item there are no significant differences across the two conditions.[18] With or without the use of a computer doctors seem to attain satisfactory standards of clinical care. In identifying complaints and background factors, conducting the physical examinations, interpreting the evidence, and reaching decisions the use of a computer seems to have had no major effect. The results for the adequacy of communication with the patient, the adequacy of records made, and efficiency in the use of time are less clear because, as the lower levels of interrater agreement show, these aspects of the consultation were less reliably measured.

Furthermore, in looking at the pattern across all 15 measured aspects of the delivery of care there is a trend towards the computer condition attaining the lower scores. In 11 of the 13 instances where there is a difference the consultations in which the computer was used are rated lower. This proportion is significantly different from that expected by chance, at $p<0·05$ level of probability (Sign test[19]). Thus, although there is no computer effect strong enough to emerge on the item by item analysis, there is evidence of a weak overall effect of computer use that is generally detrimental to the delivery of care. This does not appear to be a major concern for two reasons. Firstly, the effect is small and standards of delivery of care remain positive. Secondly, the computer system was being investigated in a trial and the doctors were relatively inexpert in its use. If such a system was used routinely such a small negative impact would probably disappear.

A criticism that might be levelled at the results is that the measures of the delivery of care were too insensitive to show the effects of computer use. This seems unlikely, however, as other analyses showed that the measures discriminated effectively between the standard of care attained by the different doctors and between the standards evident in first consultations as opposed to follow up consultations. A two way analysis of variance of the mean ratings of the delivery of care by consulting doctor and consultation type (first presentation of symptoms or follow up visit) showed a significant doctor effect on the appropriateness of the physical examination ($F=3·15$, $p=0·022$) and near significant effects on the appropriateness of the identification of background factors ($F=2·45$, $p=0·058$) and the adequacy of records made ($F=2·34$, $p=0·068$).[18] Significant effects for consultation type were shown on the appropriateness of the physical examination ($F=4·58$, $p=0·037$), the adequacy of the physical examination ($F=6·14$, $p=0·018$), and the adequacy of records made ($F=5·24$, $p=0·026$). The physical examinations were rated more highly in first consultations, and the adequacy of records was rated more highly for follow up consultations. These findings suggest that the measures of the delivery of care developed for this study were capable of showing effects of computer use during the consultation.

Discussion

The results show clearly that even when doctors are fairly inexperienced in computer use during consultations in general practice, using such technology to provide for the review and update of patients' medical history, encounter notes, and repeat prescribing information has no appreciable impact on the standard of the delivery of care. When coupled with the findings of previous research showing no major negative effects on patient attitudes this is promising for the development of computer systems for consulting rooms.[7,8] It remains to be seen, however, whether systems that require more intensive use than the one examined here—for example, those offering medical data, symptom processing facilities, or treatment protocols—are as innocuous. Moreover, as noted earlier doctors may find the use of such systems stressful and much needs to be done to develop both hardware and software that are compatible with use in the consulting room.[7]

We thank all of the people who participated in this study: the doctors, staff, and patients of the practice concerned; the rating panel, who provided their time and expertise voluntarily; Maria Platts, of the department of community medicine, who transcribed consultations, matched symptoms, and abstracted records; and our colleagues at the MRC/ESRC Social and Applied Psychology Unit, Mike Fitter, Bob Garber, and Guy Herzmark. The research was supported by the IBM (UK) Scientific Centre and the Department of Health and Social Security.

References

1 Computer Working Party. *Computers in primary care*. Occasional paper 13. London: Royal College of General Practitioners, 1980.
2 *Computing in general practice. A report for the General Medical Services Committee of the British Medical Association*. London: Scicon Consultancy International, 1980.
3 Fitter M, Herzmark G, Brownbridge G, Garber JR. *Evaluation of the Micros for GP's scheme. Interim report*. London: Department of Health and Social Security, 1984.
4 Fitter M, Brownbridge G, Garber JR, Herzmark G. A human factors evaluation of the IBM Sheffield primary care system. *Proceedings of the 1st IFIP conference on human computer interaction (Interact '84)*. Amsterdam: Elsevier Science, 1984.
5 Potter AR. Computers in general practice: the patient's voice. *J R Coll Gen Pract* 1981;31:83-5.
6 Pringle M, Robins S, Brown G. Computers in the surgery: the patient's view. *Br Med J* 1984;288:289-91.
7 Cruickshank PJ. Patient stress and the computer in the consulting room. *Soc Sci Med* 1982;26:1371-6.
8 Brownbridge G, Herzmark G, Wall T. Patient reactions to doctors' computer use in general practice consultations. *Soc Sci Med* 1985;20:47-52.

9 Herzmark G, Brownbridge G, Fitter M, Evans A. Consultation use of a computer by general practitioners. *J R Coll Gen Pract* 1984;34:649-54.
10 Donabedian A. The quality of medical care. *Science* 1978;200:856-64.
11 Brook RH. *Quality of care assessment: a comparison of five methods of peer review*. Washington DC: National Center for Health Services Research and Development, 1973. (DHEW publication No HRA-74-3100.)
12 McAuliffe WE. Studies of process-outcome correlations in medical care evaluation: a critique. *Med Care* 1978;16:907-29.
13 Rhee SO, Luke RD, Lyons TF, Payne BC. Domain of practice and the quality of physician performance. *Med Care* 1981;19:14-23.
14 Pendleton D, Schofield T, Tate P, Havelock P. *The consultation: an approach to learning and teaching*. Oxford: Oxford University Press, 1984.
15 Absolon PJ, Denner TJ, Kaye G. *The IBM Sheffield primary care computer project—introduction and objectives. Scientific report*. Winchester: IBM UK, 1983.
16 Herzmark G. Reactions of patients to video recording of consultations in general practice. *Br Med J* 1985;291:315-7.
17 Raynes NV. A preliminary study of search procedures and patient management techniques in general practice. *J R Coll Gen Pract* 1980;30:166-72.
18 Ferguson GA. *Statistical analysis in psychology and education*. Tokyo: McGraw-Hill Kogakusha, 1976.
19 Leach C. *Introduction to statistics: a non-parametric approach for the social sciences*. Chichester: John Wiley and Sons, 1979.

(*Accepted 17 July 1985*)

Audit Report

Chlorthalidone treatment in patients with hypertension associated with hypokalaemia

In an audit of the care of patients with hypertension in our practice roughly a fifth of those studied who were taking a diuretic were found to be hypokalaemic. Of 6400 patients in the practice, 354 were diagnosed as hypertensive. One hundred and fifty two of those patients, who attended one branch surgery, were investigated for the following: weight, urine analysis, electrolyte concentrations, lipid and urate concentrations, electrocardiography, chest *x* ray film, and smoking habits, and the results of blood chemistry tests were entered into a computer. Of the 152 patients, 114 were taking a diuretic as part of their treatment; 26 had potassium concentrations below 3·2 mmol(mEq)/l. Eighteen of these were taking atenolol with chlorthalidone, 13 were taking Tenoret 50, and five were taking Tenoretic. The remainder were taking one of six different agents.

The serum potassium concentrations in patients taking atenolol with chlorthalidone we've compared with those in patients taking non-selective β blockers with diuretics (sotalol, propranolol, and timolol with either bendrofluazide or cyclopenthiazide) as follows: potassium concentration <3·5 mmol/l, 29 patients taking atenolol, 22 taking a β blocker and diuretic; potassium concentration >3·5 mmol/l, nine patients and 35 patients respectively ($\chi^2=38·87$; $p<0·001$); potassium concentration <3·2 mmol/l, 18 patients taking atenolol, two taking a β blocker; and potassium concentration >3·2 mmol/l, 20 and 35 patients respectively ($\chi^2=16·88$; $p>0·001$). Although some have suggested that patients who take a diuretic do not develop appreciable hypokalaemia, Poole-Wilson suggested that mild hypokalaemia might increase the mortality in patients with hypertension because ventricular arrhythmias may be induced.[1] Patients who have hypokalaemia show an increased ventricular ectopic activity compared with a control group, which implies that such arrhythmias relate to the severity of hypokalaemia. According to Rabkin ventricular extrasystoles result in an incidence of sudden death three times greater than normal in the over 40 age group, and there is an appreciable reduction in arrhythmias after hypokalaemia is corrected.[2]—F D Skerrett, general practitioner, Drinnicks, Water Lane, Golant, Fowey PL23 1LF. (*Accepted 10 July 1985*)

1 Poole-Wilson PA. Potassium diuretics and heart muscle. In: Wester PO, ed. *Electrolytes and the heart*. New York: Trans Medica, 1983:57-62.
2 Rabkin SW. Electrocardiographic abnormalities in apparently healthy men. *Drugs* 1984;28(suppl 1):28-45.

100 YEARS AGO

The drowning season is again in full swing, and the newspaper-heading "Bathing Fatalities" threatens to remain standing for some weeks to come. A cursory analysis of the 10 or 15 cases already reported shows that, though a few of the victims were unable to swim, and unquestionably died from drowning, or, as Johnson quaintly puts it, "suffocation by water," yet a larger proportion were expert swimmers, for whose sudden loss of buoyancy and fatal submersion the popularly accepted theory of cramp is an unsatisfactory and inadequate explanation. Of this class of accidents, one case, that of a man who was drowned in the Thames on Sunday, may be quoted as typical of all. The deceased, accompanied by a friend, had swum across the stream, when the latter heard a cry, and, looking up, saw his companion throw up his arms and disappear. At the inquest it was suggested that the deceased had been seized with cramp, and the jury returned the verdict usual in such cases, a verdict which, in so far as it implies that asphyxia is the sole or even ultimate cause of death, is probably incorrect in a large number of cases. Cramp, though an exceedingly painful affection, and one to which bathers are especially liable, is quite insufficient to account for the abandonment of effort and sudden disappearance of a healthy swimmer; the very tyro in the art knows from experience, however little, that there is but a trifling difference between the specific gravity of his body and that of the fluid supporting it, and that if all his limbs be paralysed with cramp, the slightest movement is enough, even in fresh water, to keep his mouth and nose above the surface. It is true that the swimmer is tempted to over-confidence in his staying-power, and that the unaccustomed eye is apt to underestimate distance over water, so that when the self-allotted task is but half finished, its accomplishment often seems almost hopeless. But danger makes life seem dearer; comparative rest may be easily obtained by floating; and the delivery and sequence of the strokes, however feeble, is so mechanical that it would be hardly less suicidal for a disheartened or frightened swimmer to throw up his arms and voluntarily consign himself to death, than for a tired traveller to precipitate himself from the height which showed him that his destination was a few miles further off than he anticipated. It is not improbable that failure of the heart is, in many cases at least, the primary cause of death. A plunge into cold water is sufficient to throw a weak heart into a dangerous state of tension, which is further increased by swimming—an exercise perhaps the most violent in which it is possible to indulge. It is a remarkable fact that few, if any, of the subjects of so-called "cramp"—not, be it understood, the ordinary bather's cramp, but the mysterious seizure which impels its victim to throw up his arms and sink with a single cry—are ever recovered alive, and it is not an unreasonable inference that in many cases death occurs independently of, perhaps previously to, submersion. A necropsy or two might throw some light on the matter, but in the meantime it is desirable that the excellent bathing-rules of the Royal Humane Society be taught in every school, as a preliminary to instruction in swimming, and that medical men should impress upon all subjects of cardiac weakness the risks of prolonged immersion and exercise in chilly water. (*British Medical Journal* 1885;ii:265.)