PRACTICE OBSERVED

Practice Research

Comparisons between written and computerised patient histories

MARTIEN J QUAAK, R FRANS WESTERMAN, JAN H VAN BEMMEL

Abstract

Patient histories were obtained from 99 patients in three different ways: by a computerised patient interview (patient record), by the usual written interview (medical record), and by the transcribed record, which was a computerised version of the medical record. Patient complaints, diagnostic hypotheses, observer and record variations, and patients' and doctors' opinions were analysed for each record, and records were compared with the final diagnosis.

About 40% of the data in the patient record were not present in the medical record. Two thirds of the patients said that they could express all or most of their complaints in the patient record. The doctors found that the medical record expressed the main complaints better (52%) than the patient record (15%) but that diagnostic hypotheses were more certain in the patient record (38%) than in the medical one (26%). The number of diagnostic hypotheses in the patient record was about 20% higher than that in the medical record. Intraobserver agreement (51%) was better than interobserver agreement (32%), while the inter-record agreement varied from 25% (between the medical and patient records) to 35% (between the transcribed and patient records). One third of final diagnoses were seen in the medical record, with 29% and 22% for the transcribed and patient records, respectively. Interobserver agreement in the final diagnosis was 35%.

The results of the study suggest that computerised history taking is suitable for certain patients in addition to, and not as a substitute for, the oral interview with a doctor.

Departments of Medical Informatics and Medicine, Free University, Amsterdam, The Netherlands

MARTIEN J QUAAK, MD, medical research fellow
R FRANS WESTERMAN, MD, consultant physician
JAN H VAN BEMMEL, PHD, professor of medical informatics

Correspondence to: Professor Jan H van Bemmel, Department of Medical Informatics, Academic Hospital Free University, PO Box 7057, 1007 MB Amsterdam, The Netherlands.

Introduction

Diagnosis is the core of medicine; no medical decision can be made reliably without sufficient, though not necessarily complete, data interpreted by knowledgeable doctors. The most important stage in the decision scheme is the first one: the encounter between the patient and a doctor, be he a general practitioner or specialist. When medical data acquired at this stage are unreliable or deficient there is a risk of a wrong decision being made and the patient being improperly treated or referred to the wrong specialist. Taking a reliable history is the crux of all further medical actions, as was recognised by Weed when he devised the problem oriented record.[1]

Since computers made their first appearance in medicine attempts have been made to augment history taking by studies in which patient performance and acceptance were investigated.[2-4] Before interactive equipment became available many techniques were used to acquire data, with or without the help of a doctor or nurse, such as coding sheets,[5] mark sense forms, punched cards, sortable pictures or cards, audio and video tapes, etc. Later, interaction between computer and patient became possible with the use of typewriter terminals and visual display units.[6,7] Also special terminals were developed with adapted keyboards and for displaying graphical information.[8] Such systems should be highly interactive, have an ergonomic design, and contain a built in "intelligence."

Personal computers are now familiar in doctors' offices and some patients' homes and a new generation of doctors is being educated

for the coming information age.[9] More people expect doctors to use computers for practice organisation and storing patient data, and several systems have been designed to store medical records, laboratory data, diagnostic codes, and medical histories.[10]

No large studies have investigated the impact of a computerised patient history on diagnosis and treatment. Most systems emphasise the efficiency of history taking, and researchers have investigated the reactions of patients and doctors to them.[11,12] Because of the growing impact of computers on medicine the personal computer raises the possibility of automated history taking, with the medical record becoming increasingly computerised. This may lead to medical protocols and audit being based on patient data stored in computers, allowing computer assisted medical decision making using information stored in detailed data bases.

These considerations led us to investigate some aspects of computerised patient histories not hitherto reported. We evaluated how far both written and computerised medical records contained identical patient data for the same patient population; how patients reacted to automated interactive history taking; whether the diagnostic hypotheses, made on the basis of the different types of medical records, were comparable and how they related to the final diagnosis with its implications for further medical care; what interobserver and intraobserver variations occurred with these types of medical records; and how doctors reacted to the different types of medical histories.

Patients and methods

The study was carried out in the department of medicine at this hospital. Some 300 patients were asked to participate in the project, and 99 agreed. Those who refused did so because of lack of time or other obligations. All patients referred by their general practitioner had a morning outpatient appointment and were asked to come one hour earlier to participate in the project. There were 38 men and 61 women, whose mean ages were 48 and 45, respectively. There were no significant age differences between the men and women ($t=0{\cdot}88$, $p=0{\cdot}38$).

INTERACTIVE SYSTEM

We have developed an interactive system to take the computerised patient history,[13] in which the main characteristics fulfil the requirements established by earlier investigators.[14] The system consists of a display terminal and an adapted keyboard with only function keys, which is very easy to use. The terminal is connected to a small computer, which runs a program generated by fourth generation software.[15] All questions are simply formulated, and when a patient does not understand the question he may press a key to convey this. Answers are selected from a multiple choice menu with up to seven possibilities instead of just yes, no, or don't know. As soon as a key is touched the answer appears as full text on the screen so that the patient may change it, reject it, or even go back to earlier questions. Description of complaints is supported by a schematic picture of the human body, on which the patient may indicate sites of pain or discomfort.

This system consists of over 400 questions relating to 179 different items. Two hundred and sixty "help messages" were built in to help the patient when he does not understand the question. Each history taking is preceded by a brief exercise in which the patient is asked about daily habits, such as watching television, to familiarise him with the method. The system starts by asking about the patient's main complaints, which he may indicate in the picture. It then moves on to questions on related organ systems, which are screened, and in depth questioning follows only if a patient has indicated that he has complaints. When the patient has finished, or wishes to stop, a printed report is immediately generated for him to review, change, or reject. The patient receives a copy of this report.

The study consisted of three different stages.

STAGE 1: ACQUISITION OF MEDICAL RECORDS

All patients participated in the computerised history taking before being interviewed by the doctor. The computerised history taking ended in the printed report, called the patient record, and the interview with the doctor in the written medical record. This interview was taken by a medical student in the last stage of clinical training and afterwards verified by a resident. Both the computerised and the written interview followed the same guidelines and were based on the list of medical history questions routinely used in the department for patients on their first visit to the outpatient department. The computerised history was not available during the oral interview. Directly after the computerised interview the patients answered a short written questionnaire, giving opinions on this method of history taking.

As our aim was to compare the usual, written medical record with the computerised patient record a third medical record was generated by transcription of the medical record, following the same system as the patient had used, by an independent physician. This was called the transcribed record. Except for some non-quantifiable or uncodable items this record should have contained the same information as the medical record.

This phase of the project provided three types of medical records from 99 patients. A subgroup of 50 patients, drawn from and with similar age and sex distributions as the original group of 99, gave their opinion on computer supported history taking. This completed the first stage of our project, in which we compared the prevalence of complaints reported by patients both within and among the different types of records.

In order to compare the medical data in the different types of records in stage 1 we have restricted ourselves to a comparison of the quantifiable or codable data alone. It is much more difficult to compare qualitative data and impressions, let alone the results of non-verbal communication, although we are aware of the diagnostic importance of such information for the doctor. The computer compared the information contained in the so called frequency answers to certain types of questions. These answers were given by the patients in a range from "never" to "always." We clustered these categories of answers in four groups: negative (never; −), slightly positive (once and seldom; +), moderately positive (sometimes and regularly; ++), and strongly positive (often and always; +++). Depending on the frequency of a complaint, further in depth questions were asked. For such answers we identified absent, small, and large differences in the given answers for the different types of records.

STAGE 2: DIAGNOSTIC HYPOTHESES

In the second stage of the project the three types of records from a sample of 18 patients were examined in weekly batches of six by three doctors, who had not seen the patients before. All records were made unidentifiable. For six of these patients all the records were examined a second time much later without the doctors' knowledge. Each doctor was given 72 records and asked to rate their usability and indicate diagnostic hypotheses. These diagnostic hypotheses and the doctors' opinions were analysed.

STAGE 3: VARIABILITY

In the third stage the interobserver and intraobserver and interrecord and intrarecord comparisons and a comparison of final diagnoses were analysed. We looked for differences in diagnostic hypotheses between the types of records and for discrepancies between and within doctors. The three doctors were offered the complete written medical records of the same 18 patients one and a half years after their first visit to the clinic. These records were analysed for intraobserver and interobserver variations.

Statistical methods—We used the following statistical methods to test the validity of our hypotheses: Student's t test to compare the means of two groups with respect to some continuous feature; Pearson's product-moment correlation to investigate the linear relation between two continuous features; the χ^2 test to investigate the relation between two features measured with a nominal scale; and Cochran's Q test for the investigation of several nominally scaled features simultaneously. A significance level of $\alpha=0{\cdot}0$ was chosen for all tests, and continuous features were log transformed where appropriate.

Results

STAGE 1: ACQUISITION OF MEDICAL RECORDS

Frequencies of patient complaints

For each organ system we analysed the number of times that patients indicated a complaint in the written medical record and the computerised patient record in detail. The medical record was first transcribed to allow an automated analysis and comparison of the data. Here we summarise the statistics of the results for the different organ systems.

Figure 1 shows the percentage of times that a complaint was indicated as slightly (+), moderately (++), or strongly (+++) positive and the percentage of times that data were deficient. This is drawn for the

respiratory, the circulatory, and the gastrointestinal organ system; stools; the genitourinary organ system; the nervous system; general complaints and skin; various other disorders; and health disturbances.

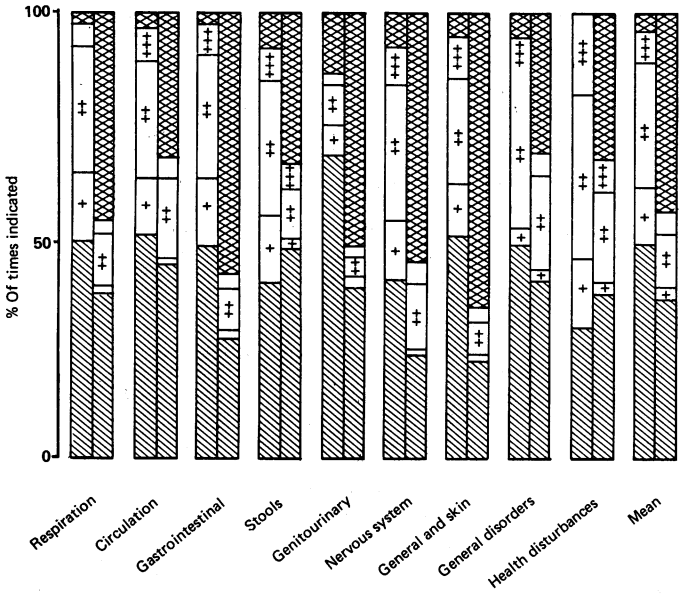

FIG 1—Frequencies of patients' complaints for the different organ systems in the written medical record after transcription (right hand column of each pair) and the computerised record (left hand column of each pair). Percentage of complaints denied by patient ▧ and complaints for which no answer was given (or asked for) ▩ are shown; + = symptoms indicated once or seldom, ++ = those indicated sometimes or regularly, and +++ = those indicated often or always.

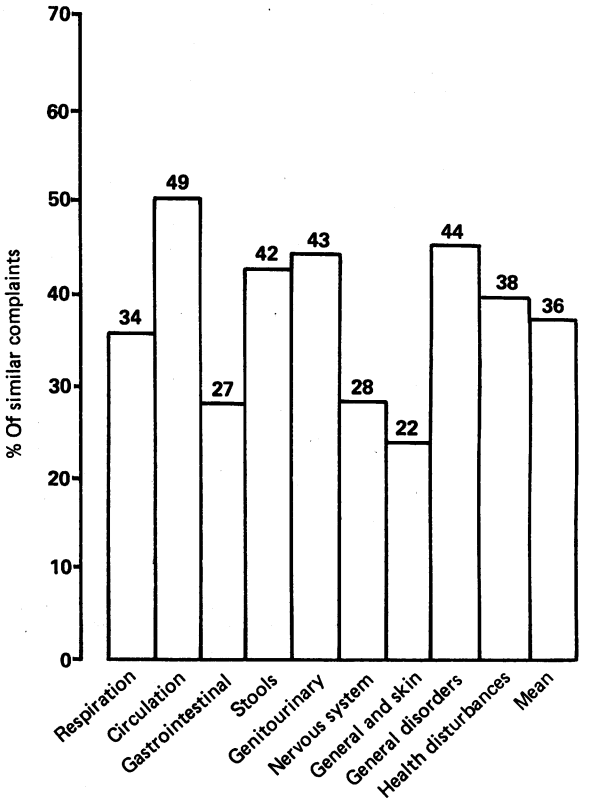

FIG 2—Complaints classified as similar in medical and patient records for all organ systems.

Many of the data (45%) were deficient in the (transcribed) written medical record but only 4% in the patient record. On average, 48% of all complaints were positively included in the patient record, compared with 19% in the medical record. The remainder, 36% for the medical record and 48% for the patient record, were negatively indicated. Of the strongly positive registrations, 8% were indicated by the patient in the patient record and 4% in the written record. Table I shows how often complaints were not present or positively indicated in the two types of records for all 99 patients. In absolute numbers the differences were largest for the moderately positive (sometimes and regularly) indications.

TABLE I—*Number of complaints reported in patient and medical records by frequency*

	Never (−)	Seldom or once (+)	Sometimes or regularly (++)	Often or always (+++)	Unknown (0)
Patient record	4158	1081	2374	747	367
Medical record	3225	159	1223	324	3765

Differences between records

Figures 2-5 show the differences between the two types of records for all organ systems. A complaint indicated with the same frequency in both records was called similar. Complaints indicated in both records but with different frequency were also coded similar but with small or large differences. A difference was called small if, for instance, the patient record showed ++ and the medical record +++, and a large difference was present between + and +++. On average, 36% of all complaints were similar, 2·6% having small and 0·5% large differences. Figure 4 shows that of 50% of the complaints indicated in the patient record, on average 14% were negated in the medical record and 36% absent. The reverse—complaints present in the medical record but not in the patient record—was true for 2·8% (negated) and 1·7% (absent) (fig 5).

Figures 2-5 show the scatter of percentages between different organ

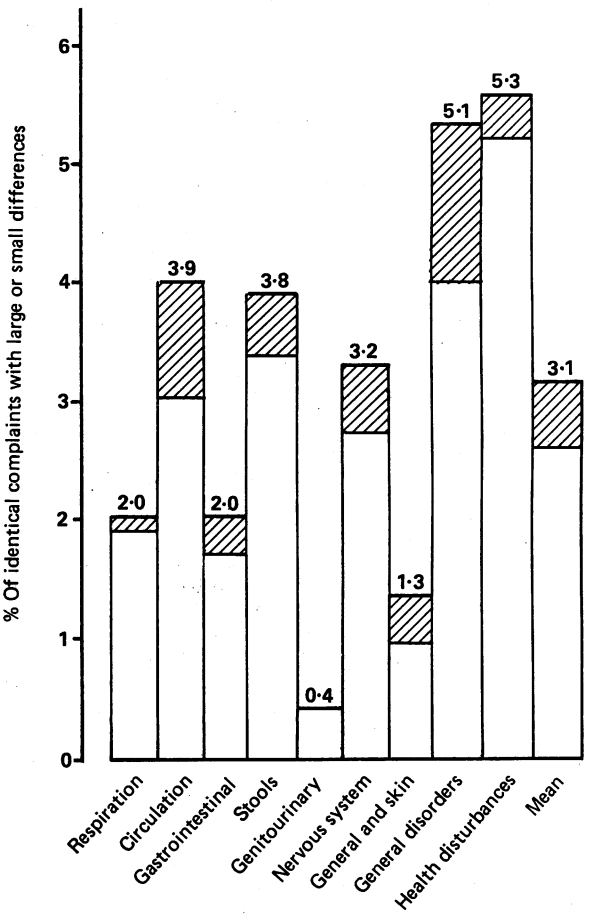

FIG 3—Complaints classified as identical in medical and patient records but with small (□) or large (▨) differences.

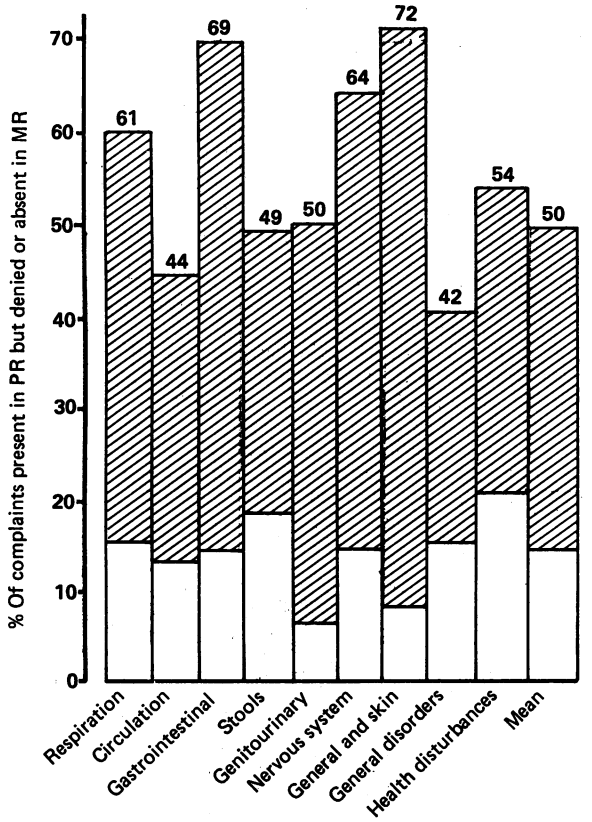

FIG 4—Complaints present in patient record (PR) but absent (▨) or denied (□) in medical record (MR).

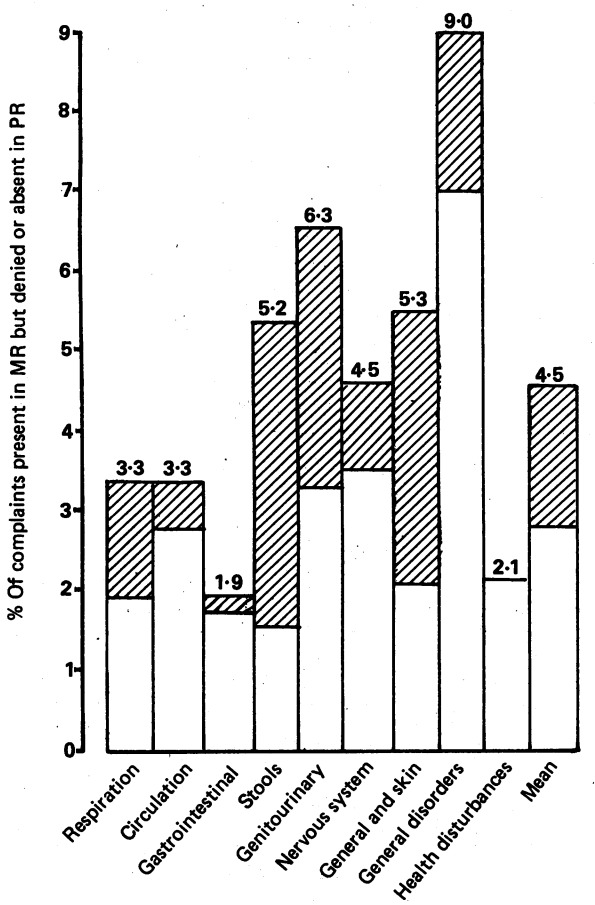

FIG 5—Complaints present in medical record (MR) but absent (▨) or denied (□) in patient record (PR).

systems. On the basis of these frequencies, no differences were more prominent in one organ system than in another. Because "strongly positive" differences seem to be more serious we will discuss some of these. In the "circulation" section patients indicated having shortness of breath on exercise several times but not during the oral interview. Patients also complained of belching or rumbling and of back pain in the computerised interview but not during oral history taking. Such differences did not seem to be very serious or frequent, but many more observations were omitted in the medical record than in the patient record.

Patients' opinions

Fifty patients answered a questionnaire on the computerised interview. The ability of patients to indicate their complaints with the help of a computerised interview was investigated (table II). Twenty per cent said that they could indicate all, 48% most, and 26% some of their complaints. Thirty six per cent of the patients could not express some physical complaints; in 22% this was their most important complaint. Patients found this method of history taking: useful 92%, easy 84%, interesting 84%, not lengthy 74%, not annoying 76%, not too difficult 72%, and not unnecessary 72%; while 10% said that their insight into their own health had changed, 78% said that it had not altered and the others gave various answers. The choice of answers was considered to be good by 76%; too restricted by 12%; and too extensive by one patient. Most (91%) found the printed report useful, 78% orderly, 58% not too long, and 62% not too short. We could find no relation between patients' opinions and age or sex (all $p>0{\cdot}05$). Further details about patients' opinions may be found in a separate report.[16]

TABLE II—*Patients' answers to questionnaire on computerised history taking. (Values are percentages of patients)*

Question	Answers	Reply
Do you think the computerised interview was (tick as many as you like)	Useful	Yes: 92
	Easy	Yes: 84
	Interesting	Yes: 84
	Lengthy	No: 74
	Annoying	No: 74
	Difficult	No: 72
	Unnecessary	No: 72
Were you able to express your complaints by the interview?	All complaints	20
	Most complaints	48
	Partly	26
	None	2
	No answer/other	4
Could all complaints be expressed by the interview? (Tick as many as you like)	All	22
	Some physical not	36
	Some psychological not	10
	Some other not	4
	Most important not	22
	No answer/other	16
Did you change your opinion regarding your own complaints while answering the interview?	Insight increased	10
	Not changed	78
	Changed	2
	No answer/other	10
What is your opinion about the range of answers in the different questions?	Too extensive	2
	Sufficient	58
	Good	18
	Too restricted	12
	No answer/other	10
What is your opinion about the printed report? (Tick as many as you like)	Useful	Yes: 91
	Orderly	Yes: 78
	Too long	No: 58
	Unnecessary	No: 53
	Unclear	No: 60
	Too short	No: 62

Patients' performance

An average of 66 minutes were needed to complete the interview. Younger patients completed the history in a significantly shorter time (within 60 minutes) than older patients. We found a significant ($p<0{\cdot}001$) correlation of $-0{\cdot}52$ between age and the number of questions answered each minute. The fast patients answered 3·5 questions a minute, the slow

patients 2·5. On average, a completed history contained 222 answers. We found no sex differences in the average time needed for completing the history (*t* test, p=0·70).

STAGE 2: DIAGNOSTIC HYPOTHESES

For a subpopulation of 18 patients all three types of records were examined by three doctors to investigate the diagnostic information contained in them. Besides these 54 records the same doctors also interpreted six repeat records of each type, making a total of 72. There was no outward difference between the transcribed record and the patient record. From the 54 records the three doctors generated a total of 522 diagnostic hypotheses: 167 from the medical records, 156 from the transcribed records, and 199 from the patient records (see fig 6). The first doctor generated 167 diagnoses in all, the second 193, and the third 162. On average, 3·3 diagnostic hypotheses were generated for each record, with 20% more for the patient records and 10% less for the transcribed records.

We also asked the doctors to characterise each diagnostic hypothesis as certain, probable, or possible. Table III shows that in the certain category the patient record was much more prominent (38%) than the medical record (25%) or the transcribed record (30%). We found significant differences in these certainty profiles among the different record types ($\chi^2=11·7$, p=0·002) and even more prominent differences among the three doctors ($\chi^2=50·2$, p<0·001).

TABLE III—*Probability of di... stic hypotheses reached by three doctors using three types of record. (Values are pe... ges)*

	Medical record	Transcribed record	Patient record
Certain	25	30	38
Probable	38	36	40
Possible	37	34	22

Doctors' opinions

The doctors were asked about the medical usability of the records and whether they thought that the records gave sufficient importance to the main complaints or contained no relevant diagnostic information. Table IV summarises the doctors' opinions (more than one answer was possible for each record). They found information about the main complaint in the medical record in 52% of the 54 records, in the transcribed record in 26%, and in the patient record in only 15%. As table IV shows, the transcribed record was considered to have least usability for all other indications, whereas the patient record was considered to be very important for organ systems related to other complaints and routine questions (both 46%). Table V amplifies these findings, showing whether the records were good or perfectly, partly, hardly, or not usable for patient care. These impressions, however, are qualitative rather than quantitative. For this question the answers were mutually exclusive. In 54% of the cases the medical record was considered to be good (39%) or perfectly usable (15%). For the patient record the figures were only 19% and 9%, respectively, and half of the patient records were considered to be partly usable. The transcribed record was considered to be not or hardly usable in 26% of cases.

TABLE IV—*Usability of three types of record for diagnosis. (Values are percentages)*

Usability (more than one answer possible)	Medical record	Transcribed record	Patient record
For main complaint	52	26	15
For organ systems related to main complaint	44	27	35
For organ systems related to other complaints	43	31	46
For routine questions	35	24	46
For general information	11	17	24
No use/unknown	11	27	2

TABLE V—*Usability of three types of record for patient care. (Values are percentages)*

Usability	Medical record	Transcribed record	Patient record
Perfect, good	54	22	28
Partly	35	39	50
Hardly, not	7	26	13
No answer	4	13	9

STAGE 3: VARIABILITY

We investigated interobserver and intraobserver variation for the doctors and for the different types of records. For the 18 patients, each with three types of records, and the three doctors 522 diagnostic hypotheses were generated. In only 17% of all of the diagnostic hypotheses was there full agreement among the doctors for all three types of records. Figure 6 shows this as a Venn diagram. Agreement between the medical and transcribed records was 32%, between the medical and patient records 25%, and between the transcribed and patient records 35%. The agreement between the first and second doctors regarding the diagnostic hypotheses was 28%, between the first and third 30%, and between the second and third 39%.

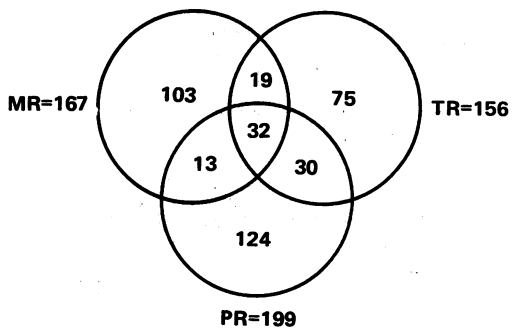

FIG 6—Venn diagram of discrepancies in diagnostic hypotheses reached by three doctors between the medical record (MR), the patient record (PR), and the transcribed record (TR). Of the 522 hypotheses, only 90 (3×30) were identical for all three doctors and 96 (3×32) for all three types of record.

We also investigated the diagnostic profiles—that is, the number of times the doctors recorded diagnostic hypotheses for the different organ systems. In this respect there were no differences between the doctors ($\chi^2=19·8$, p=0·14) but significant differences between record types ($\chi^2=30·5$, p=0·007) and sexes ($\chi^2=34·5$, p<0·001).

For six patients the three records were offered twice, resulting in 150 extra diagnostic hypotheses from 18 records and three doctors: 2·8 diagnostic hypotheses for each record. For these hypotheses we investigated the intraobserver and intrarecord variability. The intraobserver agreement varied for the three doctors from 40% to 61% and the intrarecord agreement from 50% to 52%. The overall intrarecord and intraobserver agreement was therefore 51%.

Comparisons with final diagnosis

We compared the diagnostic hypotheses with the final diagnoses, which were made by the same three doctors about one and a half years after the patients completed their treatment. The doctors were offered the complete written records (called the definite medical records) of the patients, but without the discharge diagnosis or any computerised report. On the basis of these documents they were asked again to make a final diagnosis. These diagnoses were compared with the hypotheses in the original records, and table VI shows the results. On average, the doctors made 2·5 diagnostic statements for each complete written record. The interobserver agreement was 35% (in the earlier records it was 32%). Of all the final diagnoses, 33% were already seen in the medical record, 29% in the transcribed record, and 22% in the patient record. As can be seen from table VI there were significant differences between these percentages (p<0·01, Cochran test with Q=27·7), but the outcome of this test was primarily influenced by the second doctor, and to a lesser degree by the third doctor. Overall, the doctors scored individually 26%, 29%, and 29%.

TABLE VI—*Agreements between diagnostic hypotheses in three types of record and final diagnosis for three doctors. (Values are percentages)*

Doctor	Medical record	Transcribed record	Patient record	Mean
1	25	28	25	26
2	36	30	21	29
3	39	29	19	29
Mean	33	29	22	

Discussion

Written and computerised history taking were based on the same intake questionnaire used in the department of medicine. Because we found large discrepancies between results from the written medical record and the computerised patient record in all three stages of our study we will comment on these findings. Firstly, we summarise and discuss the main differences from the three stages.

STAGE 1

In stage 1 it seemed that on average, for all patient data, 12% more negative answers were given in the patient records than in the medical ones and that about 40% of data present in the patient records were not observed in the medical records. This finding is supported by several other studies.[3] In the category of complaints that were always or often present patients gave twice as many indications in the patient records as in the medical ones. The same was true for indications in the categories sometimes or regularly. For borderline answers (seldom, once) the discrepancies were even larger. In cases where patients indicated a complaint in both records discrepancies were minor: large differences in 0·5%, small ones in 2·6%, and fully similar indications in 36%. A large percentage was denied or not present in the medical record but indicated in the patient record: 14% and 36%, respectively. The reverse—present in the medical record but denied or not given in the patient record—was true in 2·8% and 1·7% of cases, respectively. While 68% of the patients said that they could give most of their complaints in the patient record, 22% said that they could not mention their main complaint.

The lack of data in the medical record seemed to be spread equally over all the organ systems included in the history. We should, however, be careful not to conclude too hastily that data present in the patient record were indeed missing from the medical one. With most computerised interview systems the patient is stimulated to give more answers because more information is requested so that some redundancy of data seems inevitable. With our system, however, some "intelligence" has been built into the branching logic, and questions are asked in more depth only if the patient has indicated some complaints. Our study confirms other findings that patients seem to be very positive about computerised history taking.[3 7 11 12]

STAGE 2

The findings in stage 2 concur with those in stage 1. The doctors, however, thought that main complaints were much better expressed in the medical record (52%) than in the patient record (15%) and that there was also more information about the organ systems related to the main complaint in the medical record: 44% compared with 35%. Other organ systems and routine questions, however, were considered to be better represented in the patient record (46% and 46%, respectively, whereas for the medical record the figures were 43% and 35%). This is further supported by the opinions on usability: the medical record was considered to be good or perfectly usable for patient care in 54% of cases and the patient record in 28%; the medical record was described as partly usable in 35%, and the patient record in 50%. The transcribed record was valued below the other records, probably because of the filtering process of the transcription, which meant that non-factual and non-quantifiable data could not be documented.

In general, the patient record was considered to contain hard facts and to be more complete and reliable. Interestingly, diagnostic hypotheses for the patient record were labelled as certain or probable in 78% of cases, while for the medical record this was only 63%. The absolute number of diagnostic hypotheses generated for the 18 patients for the patient record was about 20% higher than for the medical one, reflecting the larger amount of data in the patient record.

STAGE 3

Stage 3 showed large discrepancies between the doctors, irrespective of the type of record. The average agreement between two doctors was only 32%, while the mean intraobserver agreement was 51%, highest for the medical record (61%) and lowest for the transcribed record (36%). The interrecord agreement varied from 25% (between the medical and patient records) to 35% (between the transcribed and patient records). These interobserver and intraobserver and interrecord variabilities shed more light on the usability of the different types of records.

We have shown that 33% of the final diagnoses were reflected in the initial diagnostic hypotheses derived from the medical record. For the transcribed and patient records this was even less: 29% and 22%, respectively. Because the medical and transcribed records seemed to agree more with the definite medical record than the patient record it could not have been merely the printed format that caused this difference. The medical record seems to contain better semantic information, whereas the patient record contains more factual data, not necessarily leading to diagnostically relevant conclusions. We should be aware, however, that this is not a final conclusion because on the basis of these and similar findings we should improve computer assisted history taking, which is only in its infancy. Furthermore, it should be realised, as has been shown by Hampton *et al*,[17] that for medical outpatients the clinical examination rarely adds to the history diagnostically. Our study confirmed this, and we were able to extend this finding to the three types of record.

ROLE OF COMPUTERS

The fact that the doctors did not find the same diagnostic information in the patient or transcribed records as in the medical record could be attributed to several reasons. Firstly, all of the doctors are used to written records. The batch of six records that they received each week contained four computerised records (two patient and two transcribed records). As the doctors saw many more written records during the week they could not be expected to adapt to this uncommon presentation of patient history data. Moreover, we should not exclude the fact that many doctors are still slightly prejudiced against computerised patient histories. The most important reason, however, is the fact that computerised records contain only the formal, factual aspects of a history. Qualitative and non-verbal information, let alone personal notes, are not contained in such structured records, nor are they written in natural language.

The large discrepancy between the doctors, even though they were all from the same department, deserves comment and might have several explanations. The diagnostic statements that were requested from the doctors were based on the written or printed records alone; the doctors did not see the patients themselves. Nevertheless, the doctors felt significantly more certain about hypotheses drawn from the patient record, perhaps because it contained more data, as shown in stage 1 and supported by stage 2.

CONCLUSIONS

From these findings several conclusions may be drawn.

Firstly, computerised patient histories are more complete than

written medical records. This does not necessarily lead to more diagnostically important conclusions. Our study, however, shows that there were more than twice as many strongly positive complaints in the computerised patient record than in the written medical record. This is supported by the fact that 68% of all patients said that they could express all complaints; most (92%) were positive about computerised history taking.

Secondly, doctors generated about 20% more diagnostic hypotheses for the patient record than for the medical one, with an average of 3·3 hypotheses for each record. Though the doctors preferred the medical record to the patient one (54% *v* 28% when asked about usability), they were much more certain about diagnostic hypotheses in the patient record (38%) than in the medical one (25%).

Thirdly, interobserver agreement in interpreting medical records seems—rather surprisingly—to be independent of the type of record and was as low as 32%; even for the final medical record it was only 35%. Intraobserver agreement also seems independent of the record type (51%). This is clearly an area of concern, and further research is needed. Widespread acceptance of computerised history taking cannot be hoped for if such large variabilities persist.

Fourthly, 33% of the diagnostic hypotheses from the written medical record concurred with the final diagnosis, while for the computerised patient record the figure was only 22%. This is only partly due to the formal character and the structure of a computerised record, as was made clear by the comparison with the transcribed record. The main cause probably lies in the smaller amount of semantic information contained in computerised records, and thus the interpretation of both written and computerised medical records needs improvement.

Finally, our study suggests that computerised history taking is suitable for certain patients (first referral, chronic diseases, and follow up), preceding rather than replacing the oral interview in order not to miss any relevant data for further diagnosis and treatment.

We thank the consultant physicians J A Schouten and L van Bergeijk for their cooperation and help in the interpretation of the medical records; P J G van der Voort and E J P M de Moel for the data handling; and Professor F M Hull and J Kuik for their valuable suggestions.

References

1 Weed L. *Medical records, medical education and patient care*. Cleveland: Case Western Reserve University Press, 1971.
2 Coombs GJ, Murray WR, Krahn DW. Automated medical histories: factors determining patient performance. *Comput Biomed Res* 1970;3:178-81.
3 Grossman JH, Barnett GO, McGuire MT, Swedlow DB. Evaluation of computer-acquired patient histories. *JAMA* 1971;215:1286-91.
4 Slack WV, Slack CW. Patient computer dialogue. *N Engl J Med* 1972;286:1304-9.
5 Mellner C, Selander H, Wolodarski J. The computerized problem oriented medical record at Karolinska Hospital; format, function, users acceptance, patient attitude to questionnaires. *Methods Inf Med* 1976;15:11-20.
6 Card WI, Lucas RW. Computer interrogation in medical practice. *International Journal of Man-Machine Studies* 1981;14:49-57.
7 Trell E. Interactive computer program for self-distributed medical questionnaires: notes on technical implementation and use. *Med Inf (Lond)* 1983;8:139-43.
8 Somerville S, Evans CR, Pobgee PJ, Bevan NS. MICKIE—experiences in taking histories from patients using a microprocessor. In: Barber B, ed. *Medical informatics Europe*. Munich: Springer Verlag, 1979:713-22.
9 Physicians for the twenty first century. *The GPEP report*. Washington: The Association of American Medical Colleges, 1984.
10 Barnett GO. COSTAR, a computer based medical information system for ambulatory care. *Proceedings of the Institute of Electrical and Electronics Engineers* 1979;67:1226-37.
11 Logie AR, Madirazza JA, Webster IW. Patient evaluation of a computerized questionnaire. *Comput Biomed Res* 1976;9:169-76.
12 Pringle M, Robins S, Brown G. Computer-assisted screening: effect on the patient and his consultation. *Br Med J* 1985;290:1709-12.
13 Quaak MJ, Van der Voort PJG, Westerman RF, Hasman A, Van Bemmel JH. Automation of the patient history, evaluation of ergonomic aspects. *Int J Biomed Comput* (in press).
14 Collen MF. Patient data acquisition. *Med Instrum* 1978;4:222-5.
15 Quaak MJ, Van der Voort PJG, Van Bemmel JH. AIDA for the automation of the patient history. *Methods and Programs in Biomedicine* (in press).
16 Quaak MJ, Westerman RF, Schouten JA, Hasman A, Van Bemmel JH. Patient appreciations of computerised medical interviews. *Med Inf (Lond)* 1986;11:339-50.
17 Hampton JR, Harrison MJG, Mitchell JRA, Prichard JS, Seymour C. Relative contributions of history taking, physical examination and laboratory investigation to diagnosis and management of medical outpatients. *Br Med J* 1975;ii:486-9.

(Accepted 2 April 1987)

DOCTORS IN SCIENCE AND SOCIETY

The makings of an editor

Thomas Wakley passed the examination of the Royal College of Surgeons in 1817, at the age of 22. He married well and settled down to practise as a surgeon in Argyll Street. Here he might have had a successful if undistinguished professional career had it not been for the events of 1820. In January of that year the old mad King died. The following month a group of radical desperados, intent on murdering the Prime Minister and his entire cabinet, were apprehended. The five ringleaders of what came to be known as the Cato Street Conspiracy were duly hanged on May Day 1820 outside Newgate Prison. As their bodies were cut down, a figure dressed in sailor's clothes and with face masked appeared on the scaffold and skilfully decapitated the corpses. Obviously an expert, he was in fact Tom Parker, anatomy assistant at St Thomas's. Rumour put it about, however, that the masked man had been a surgeon from Argyll Street, and the only surgeon living there was Wakley. In August a gang of men supposedly sympathetic to the Cato Street conspirators burst in on Wakley, assaulted him, and burnt his house to the ground. Subsequently the unfortunate Wakley was accused of having been his own arsonist to obtain the insurance, a calumny he successfully contested in court with the insurance company, but he had lost his house and his practice, a disaster to a young man within six months of his marriage.

It was at this time that Wakley met William Cobbett, the radical reforming journalist, then editor of the *Weekly Political Register* and the *Evening Post*. Cobbett had exhumed the bones of Tom Paine and preserved them in his home. He had had some experience of attacking the medical establishment during an earlier part of his life in the United States. In his periodical *The Rush Light* he had for two years "flung the worst abuse that any honest physician had to bear" at Dr Benjamin Rush of Philadelphia, violently attacking the murderous regimen of bleeding and purging for which he was famous. Cobbett, like Wakley, believed himself to be a target of the Cato Street conspirators' friends and this was the bond that brought the two men together. Cobbett, as ardent a supporter of political reform as Wakley was to be for the reform of the medical profession, played an important role in encouraging Wakley to take up radical medical journalism.—CHRISTOPHER BOOTH.

Adapt and reproduce

For the future, however, the most important issue is the survival of clinical research. As biologists we can perhaps take heart from evolution, and as scientists we are incurably optimistic. Clinical research is a tender and delicate organism requiring constant care and attention, but as with all other species two attributes are vital: the capacity to adapt to changing circumstances and the ability to reproduce. I believe that clinical science has demonstrated remarkable adaptability. It has already adapted to cuts—which are not always a total disaster, since they can enable some effective pruning to be done, and there are few universities or research institutes that do not have some dead wood. It is also increasingly adapting to the new science.

But is the clinical research community reproducing itself? Here the most important factor is the recruitment and support of the young. We need, as Sir John McMichael once put it in a memorable phrase, to have the young upon our shoulders, not trample them under our feet. It is for their elders to lead by example and for their teachers constantly to encourage a questing frame of mind, something at which Sir George Pickering so greatly excelled. But in the final analysis if you want to encourage clinical research and you are a professor or a research director what do you actually do? You pick a good man at as young an age as you can, you give him all the support he needs, and you let him get on with it.—CHRISTOPHER BOOTH.