with menorrhagia, while Basu¹³ has shown that there is usually an increase during menstruation of fibrinogen degradation products in the general circulation, which is excessive in women with menorrhagia. This increase is not due to increased fibrinolytic activity in the systemic blood,¹⁴ but probably to an increased fibrinolytic activity within the uterus itself.¹⁴ The precise relationship between menorrhagia and increased fibrinolytic activity in the uterus is not certain, and both may have a common cause, such as increased vascularity of the uterus or altered activity of the steroid hormones.

Whatever the mechanism may be there is now rather more than a purely empirical basis for using antifibrinolytic drugs in treating dysfunctional menorrhagia. So far there has been no tie up between hormone levels and the increased fibrinolytic state in menorrhagia, and though hormones are still a more effective form of treatment of this, the antifibrinolytics provide a second line when hormone therapy is contraindicated. Unfortunately they too have the disadvantage of causing possible thromboembolism.

(This article will be continued in next week's *B.M.J.*)

- ¹ Corker, C. S., and Exley, D., *Steroids*, 1970, **15**, 469.
- ² Corker, C. S., and Naftolin, F., *Proceedings of the Royal Society of Medicine*, 1970, **63**, 1047.
- ³ Markee, J. E., *Progress of Gynaecology*, 1946.
- ⁴ Salvatore, C. A., *Journal of the International College of Surgeons*, 1958, **29**, 599.
- ⁵ Green, J. D., and Harris, G. W., *Journal of Endocrinology*, 1947, **5**, 136.
- ⁶ McCann, S. M., Taleisnik, S., and Friedman, H. M., *Proceedings of the Society for Experimental Biology and Medicine*, 1960, **104**, 432.
- ⁷ Harris, G. W., *In Control of Ovulation*, ed. C. A. Villee, p. 56, Oxford, Pergamon Press, 1961.
- ⁸ Gregory, N., Walpole, A. L., Charlton, H. M., and Reed, M., in *Pharmacology of Hormonal Polypeptides and Proteins*, ed. L. Martini and R. Paoletti. New York, Plenum Press, 1968.
- ⁹ Macourt, D., *Oxford Medical School Gazette*, 1971, **23**, No. 2, p. 44.
- ¹⁰ Ryan, K. J., *Acta Endocrinologica*, 1963, **44**, 81.
- ¹¹ Mikhail, G., *Clinical Obstetrics and Gynecology*, 1967, **10**, 29.
- ¹² Rybo, G., *Acta Obstetrica et Gynecologica Scandinavica*, 1966, **45**, 429.
- ¹³ Basu, K. H., *British Medical Journal*, 1970, **1**, 74.
- ¹⁴ Beller, F. K., Goebelsmann, U., Douglas, G. W., and Johnson, A., *Obstetrics and Gynecology*, 1964, **23**, 12.

Computers in Medicine

Simulation of Clinical Diagnosis: A Comparative Study

F. T. DE DOMBAL, JANE C. HORROCKS, J. R. STANILAND, P. W. GILL

British Medical Journal, 1971, **2**, 575-577

Summary

This paper presents a comparison between three different modes of simulation of the diagnostic process—a computer-based system, a verbal mode, and a further mode in which cards were selected from a large board. A total of 34 subjects worked through a series of 444 diagnostic simulations.

The verbal mode was found to be most enjoyable and realistic. At the board, considerable amounts of extra irrelevant data were selected. At the computer, the users asked the same questions every time, whether or not they were relevant to the particular diagnosis. They also found the teletype distracting, noisy, and slow.

The need for an acceptable simulation system remains, and at present our Minisim and verbal modes are proving useful in training junior clinical students. Future simulators should be flexible, economical, and acceptably realistic—and to us this latter criterion implies the two-way use of speech. We are currently developing and testing such a system.

Introduction

Two years ago a preliminary study drew attention to an increasing shortage both of clinical teachers and of patients suitable

for teaching and reported the initial results of attempts to simulate the diagnostic process by using a computer-based system.¹ Several important questions, however, were left unanswered, for though simulation seemed to offer certain advantages in teaching clinical diagnosis it was far from clear whether the computer had to be used in an on-line real-time situation (or even whether the diagnostic simulator need be computer-based at all). To investigate the situation further we have therefore constructed several simulators of varying types, and these have been studied in the department of surgery since the summer of 1970. We report here a critical analysis of our experience to date.

Modes of Simulation

Computer-based Simulator.—This mode was derived from the computer-based diagnostic simulator previously described in detail^{1, 2} but with two major changes from the system then reported. Firstly, a Computer Technology Limited Modular One computer replaced the Elliott 903C computer originally used, and the Modular One computer was used within a multi-access format. Secondly, since the initial teaching programme exerted considerable control over the student's data-gathering processes, for the present study we had to write a completely new set of teaching programmes, which differed from their original counterparts in that they allowed the user to collect clinical information in whatever quantity and whatever sequence seemed most appropriate. These were written by two of us (J.C.H. and F.T. de D.) during the first half of 1970.

Minisim Mode.—The second and third modes of simulation did not involve the use of computer-based systems. The "Minisim" mode was derived in part from work carried out by Rimoldi and associates at the Loyola Institute of Psychometrics in Chicago³ and consisted of a large board on which were

University Department of Surgery, General Infirmary, Leeds

F. T. DE DOMBAL, M.D., F.R.C.S., Senior Lecturer in Surgery

JANE C. HORROCKS, Assistant in Programming

J. R. STANILAND, Assistant in Surgical Research

P. W. GILL, M.B., CH.B., Research Assistant (Present Address: Welsh

National School of Medicine, Royal Infirmary, Cardiff.)

hung a series of cards. Each of the cards represented a patient attribute, so that the whole board might be said to represent a "case history." The "patient" was interrogated by picking cards from the board, one by one, relevant clinical information being printed on the reverse side of each card.

Verbal Mode.—In the third simulator mode patient information was recorded on a series of cards as before, but instead of these being displayed on a board they were contained within a serially-numbered box. To obtain patient information the user asked questions verbally as though the operator were the patient, and the patient's answers were provided on cards as in the previous mode.

Case Material

A total of 18 case histories were obtained from randomly selected real-life patients with abdominal pain, either on the wards or in the outpatient department (Table I). To ensure that the groups of patients in each mode were of comparable diagnostic "difficulty" levels, the 18 patients were graded according to ease of diagnosis by a number of independent clinicians. Cases 1-6 were graded as "easy," 7-12 as "moderate," and 13-18 as "difficult" diagnoses. Subsequently, we ensured that each simulator mode contained two easy, two moderate, and two difficult cases. Within this framework the allocation of cases to each simulator mode was randomized.

TABLE 1—Final (Operative and Pathological) Diagnoses in 18 Real-life Cases used for Simulation

Cholecystitis	4 cases
Diverticulitis	3 cases
Appendicitis	3 cases
Perforated peptic ulcer	2 cases
Ectopic pregnancy	2 cases
Pancreatitis	1 case
Carcinoma colon	1 case
Small-bowel obstruction	1 case
Mesenteric adenitis	1 case

Methods of Study

Validation Phase.—Before any attempt was made to evaluate the various simulator modes we undertook a "validation" phase in which clinicians and students worked through a varying number of cases in each mode to validate the computer programmes and to provide basic operational experience in each simulator mode. In all a total of nine clinicians and 10 non-clinicians (medical students and ancillary staff) worked through a total of 204 diagnostic simulations. The results, though interesting in themselves, were discarded, since the simulator modes were altered slightly in the light of experience gained during the period of this phase.

Evaluation Phase.—This phase aimed to conduct a further experiment under more controlled conditions. A group of 15 clinicians was selected (none of whom had participated in the earlier validation trials), and they were asked to work through six "cases" in each mode, the order in which the modes were presented to each clinician being randomized in advance. In practice, problems were encountered with the use of the computer-based mode in that its availability was restricted and its reliability varied from time to time. Thus only 10 out of the 15 clinicians could work through the cases stored within the computer, since we did not feel justified in attempting to assess the computer-based simulation system at times when it was not working reliably. Nevertheless, in this phase the 15 clinicians were eventually able to work through a total of 240 simulated diagnoses.

Findings

We have attempted to analyse our findings under two main headings—namely, (a) comparisons of the clinicians' performances at each of the various diagnostic simulators, and (b)

the clinicians' reactions to each of the simulators which they were asked to use.

PERFORMANCE

Some details of the clinicians' performances are shown in Figs. 1 and 2 and Table II. Several interesting variables are measured—the mean number of questions asked by each clinician, Kendall's coefficient "W", and also the percentage of accurate diagnoses. For comparison some preliminary data are shown indicating the performance of a comparable group of clinicians from our real-life studies.⁴

FIG. 1—Number of questions asked in each mode of simulation.

FIG. 2—Kendall's coefficient "W" indicating similarity between each clinician's six diagnostic performances (see text). Comparison between three modes of simulation, showing inflexible approach to computer-based system. Values indicate means and standard errors for 15 clinicians' performances on Minisim and verbal modes, and 10 clinicians at the computer-based system. For these 10 clinicians W is significantly higher when using the computer ($P < 0.01$).

TABLE II—Accuracy of Diagnosis in Various Simulation Modes

Mode	No. of Clinicians	No. of Cases	No. Correctly Diagnosed	% Correctly Diagnosed
Minisim ..	15	90	66	73.3
Verbal ..	15	90	65	72.2
Computer ..	10	60	37	61.7
Total ..	15	240	168	70.0

With the number of questions (Fig. 1) clinicians behaved in the verbal mode much as in real life. When faced with the Minisim board, however, they behaved in a totally different fashion, "picking off" extra items of information so that they finished with almost twice as many patient data as in any other mode. At the computer (even with typing assistance provided by us) the number of questions asked was often slightly fewer than in real life.

Kendall's coefficient W (Fig. 2) may be regarded as measuring the "concordance" or similarity between the clinician's six diagnostic performances in each simulator mode. If he adopted a rigid inflexible approach, asking the same questions in the same sequence for each simulated diagnosis, this was reflected in a high level for W; conversely, if he approached the cases flexibly W was correspondingly lower. With the verbal and

Minisim modes the values for W approximate to the upper end of the real life range; but at the computer-based simulator the clinicians adopted a quite different and more rigid approach, asking the same questions each time (and generally in the same sequence), despite being presented with a series of widely different diagnostic problems.

For accuracy of diagnosis (Table II) there was little to choose between the modes, particularly between the Minisim and verbal modes. The computer-based mode resulted in a slightly lower mean score, but this difference is not statistically significant ($P > 0.1$).

USER'S REACTIONS

Any simulation system should be both acceptable and enjoyable if optimal use is to be made of it. We asked the users of the various simulator modes to assess them in two ways, both as regards enjoyment of the mode and as regards its apparent realism. In both instances the result was the same. The users' preference on both counts—almost unanimously—was for the verbal mode, with the Minisim board second and the computer-based system third.

Discussion

Much caution is necessary in interpreting these results, since (for reasons already mentioned) we have not been able to carry out as formal a controlled evaluation of our three simulator modes as we could have wished, and this applies particularly to the use of the computer. Nevertheless, two accepted principles of simulation are (a) that a simulator should induce behaviour similar to that in real life, and (b) that subjects should accept the replica and be motivated to use it;⁵ on both counts the data presented here enable some tentative conclusions to be drawn.

MINISIM BOARD

Rimoldi³ evaluated a similar device to the Minisim board which he described as "the test of diagnostic skill." Our own data make it clear that his use of the definite article is hardly justified since clinicians behaved very differently from real life, specifically by picking off numerous questions which were often quite irrelevant to the case in question. Nonetheless, the board is inexpensive and it has some advantages, particularly for very junior students. Potential questions are displayed without need for the student to formulate them verbally, a useful point in early training; and the Minisim board can be used by students on a self-service basis. (We ourselves are at present using it by leaving out the board with a "case of the week" displayed. When the student has finished collecting cards and made his diagnosis he exchanges these for a data sheet showing the correct diagnosis together with the reactions of various clinicians to the same diagnostic problem.)

COMPUTER-BASED SYSTEM

Here we could not make as thorough an evaluation as we would have liked. Nonetheless, our results were far from encouraging, in that the clinicians behaved in a rigid inflexible manner, asking the same questions in every simulation, irrespective of their relevance to the particular case. Usually when asked why this was so clinicians answered that at least they knew that the computer could answer that question, and preferred an irrelevant answer to none at all.

It seems unusual that clinicians should attempt to diagnose six widely different diseases (appendicitis, cholecystitis, carcinoma of the large bowel, diverticulitis, small-bowel obstruction, and mesenteric adenitis) by a virtually identical interrogation of each case, and it is scarcely surprising that

their diagnostic accuracy was relatively low when this mode was used. Other discouraging features of the computer-based system were the poor ratings as regards enjoyment and realism, accuracy of diagnosis, and the relative costs of the systems involved.

We took good care only to make use of the computer-based system when it was working satisfactorily, even though this meant the loss of many hours of potentially available computer time. This also meant that we were unable to study our clinicians in depth to determine whether continued usage would reverse the trends, though what information we have suggests that this is not the case.

Moreover, in this study clinicians were given considerable help in using the system. We attempted to carry out parallel studies without this help but were unable to do these in more than a fraction of the clinicians. However, the results roughly confirmed the trends already noted; in fact, if anything they were more pronounced and, as Anderson⁷ has cogently remarked, typing in a "conversational" mode is not something which comes easily—either to clinicians or to medical students. Finally, though we used clinicians rather than students in our formal trials student experiences closely paralleled those of the clinicians.

CURRENT DEVELOPMENTS

The need for simulation of the diagnostic process remains. The shortages mentioned in the previous report (as regards patients and medical teachers) are becoming more acute. Indeed, our experience to date of several hundred simulated diagnoses has convinced us that such teaching techniques are well worth further experiment and development and that simulation of the diagnostic process has a future both in teaching and in research. But we find it increasingly difficult to believe that the computer can adequately sustain its central role as was originally envisaged, for the problems of real-time communication with a computer in a conversational mode are considerable. We believe that the criteria by which future simulators must be judged are economy, flexibility, and realism—and to us this last criterion implies the use of the spoken word.

At present our verbal and Minisim modes are proving of value in introducing junior students to the problems of clinical diagnosis. Nevertheless, there is clearly room for further development, and ideally the student should speak to the simulated "patient" and hear the "patient's" spoken reply in real time. We are currently in the process of constructing such a "third-generation" simulator—which makes use of another computer-based system but in a different, supportive role—and we hope to report a critical evaluation of these further developments.

We are most grateful to Professor J. C. Goligher for his encouragement and advice throughout our series of studies and also for permission to study patients under his care. We are grateful also to the Research Council's Computer-based Learning Group for the use of their Modular One computer. We are particularly pleased to acknowledge our debt to the various clinicians whose experiences and impressions form the basis of this study, and without whose active participation this work would have been quite impossible. Finally, two of us (J.C.H. and P.W.G.) were aided by a grant from the Medical Research Council, which we also gratefully acknowledge.

References

- ¹ de Dombal, F. T., Hartley, J. R., and Sleeman, D. H., *Lancet*, 1969, 1, 145.
- ² de Dombal, F. T., Hartley, J. R., and Sleeman, D. H., *British Journal of Surgery*, 1969, 56, 754.
- ³ Rimoldi, H. J. A., in *The Diagnostic Process*, ed. J. A. Jacquez, p. 315. Michigan, Malloy Lithographing, 1964.
- ⁴ de Dombal, F. T., and Gill, P. W., unpublished data, 1971.
- ⁵ Fraser, T. M., *Philosophy of Simulation in a Man-Machine Space Mission System*, U.S.-N.A.S.A. SP-102. Washington, D.C., Scientific and Technical Information Division, N.A.S.A., 1966.
- ⁶ Rimoldi, H. J. A., *Journal of Medical Education*, 1961, 36, 73.
- ⁷ Anderson, J., personal communication, 1970.