

A diagnostic support system in general practice: is it feasible?

Jacobus Ridderikhoff *, Egbert van Herk

Department of Family Medicine, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, Netherlands

Received 31 October 1996; received in revised form 3 April 1997; accepted 8 April 1997

Abstract

A medical diagnostic decision support system (DDSS) has been developed for and tested in general practice. Two major issues have been addressed: diagnostic support and usefulness. The diagnostic support pertains to the ability of the system to generate diagnostic hypotheses from a set of patient data. The usefulness is approached by creating a computer system which can be used simultaneously with the doctor–patient consultation. The support function operates by matching symptoms from the patient data base with symptom configurations contained in the knowledge base. The support is presented as a list of diagnostic hypotheses ranked by degree of concordance. A user-friendly interface has been constructed with a comprehensive set of clinical terms within which the doctor can locate a desired symptom and store it with a single keystroke. With another keystroke the doctor can check the stored data and ask for support at any moment during the process. The overall purpose is to invite the doctor to rethink and re-examine his steps and to reconsider possible alternatives in the light of the presented diagnostic information. In our view it has to be the doctor who makes the final judgement. A test with the system in general practice revealed good performance of the system and an astonishing proficiency of the participating doctors in its use during the consultation. Twenty doctors solved five patient cases, entering 2000 clinical items within acceptable limits of consultation time. In 96% of the cases the correct diagnosis appeared in the differential diagnosis list. The doctors' diagnostic accuracy was 43%. The use of standardised terminology as an option for further development is discussed. The role of the doctor in computer-aided diagnostics remains open to debate. A computer-aided diagnostic support system in general practice appears to be feasible. © 1997 Elsevier Science Ireland Ltd.

Keywords: Computers; Diagnosis; Decision support; General practice; Performance in practice

1. Introduction

It is widely assumed that general practice is about the last place where a medical decision support system can be developed and em-

* Corresponding author. Tel.: +31 10 4528069; fax: +31 10 4360717.

ployed. The structural blend of widely varying problems, whether they are medical, psychological or social, seem to make the invention and implementation of such a system rather illusory. It is, however, in this environment that the need for a decision support system is most pressing. Often the general practitioner (GP) has to act on the basis of just a few facts, vague ideas and uncertain prognoses. Furthermore, most diseases in their early stages present in a way that sometimes diverges essentially from what we may find in medical textbooks. Many GPs get bogged down in an uncertain search for cues and often come to some sort of conclusion on the basis of just a few hints. In these cases some (diagnostic) help would be more than welcome.

Systems which cover the broad spectrum of illnesses as seen in general practice, however, do not exist. Most medical diagnostic decision support systems have been developed 'in carefully defined, artificially narrow domains' [1]. A recent study of the functioning of four available computer-based diagnostic systems gave disappointing results in terms of diagnostic support [2]. The correct diagnosis came up only half to three-quarters of the time; two computer programs for acute abdominal pain scored approximately 60% [3] and Elstein et al. [4] found only 38% of correct diagnoses in a list of DSS generated diagnostic possibilities.

In routine clinical practice these kinds of system do not seem an adequate tool. They are often time-consuming and difficult to operate. Daily use is a precondition to give the practising physician sufficient proficiency to actually use the system. The system should act as a ready-at-hand consultant; a consultant who may be asked any time the GP feels a need for advice and feedback. But this thought assumes mutual understanding between doctor and computer: e.g. speaking the

same language, following the same line of reasoning, giving insight into possible errors and proposing relevant diagnostic suggestions. It is especially in these areas that the limitations of most existing computer programs surface. Among these limitations standardisation of terminology is one of the stumbling blocks [5]. Another is a user-unfriendly interface often requiring complicated and time-consuming man-machine interactions.

Starting from these thoughts we have tried to design a medical diagnostic decision support system (DDSS) which might become accepted as a welcome partner in general practice [6]. Whether our ideas could be converted into a working program was the big challenge. The other question was whether the program would operate in the setting for which it was meant. We shall elaborate on these matters in the next sections.

2. System design

From an investigation among general practitioners four prime requirements could be distilled.

1. It should be used simultaneously with the patient-doctor encounter: it should not interfere with the consultation and not detract from the patient-doctor contact.
2. It should be a logical extension of the problem-solving process of the doctor, not a replacement of it.
3. It should include a full administrative system, replacing the paper record and obviating the need for it.
4. It should present a list of diagnostic possibilities from the collected data pertaining to the case.

From this we formulated the following conditions:

the system should cover broad areas of medical diagnostics;

the system has to overcome the problems of different terminology not only between doctors but also between doctor and patient;

the system should provide the possibility to detect and flag errors in reasoning and judgement.

Much emphasis has been placed on the necessity to obtain accurate data. Although this may seem self-evident, most doctors have their own opinions about accuracy and even in cases where consensus exists they may make use of different vocabularies. Sutton [7] found that structuring and standardisation alone, even without a computer, led to a considerable improvement of diagnostic accuracy. It is obvious that computer programs can only perform adequately if the entered data are accurate and match the terminology of the data in the system's knowledge base. Elstein et al. [4] report that it is often impossible for a physician to convey a complex understanding of a case to a computer program. The more structured and the more comprehensive a patient database, the better the performance of the support system may be expected. We therefore started from the principle: 'use a predefined terminology rather than each doctor's own.' This means that the doctor has to find an item within the predefined data system which reflects as closely as possible the contents and meaning of the patient's complaint.

To that end we developed the system coding system (SCS), a hierarchically structured framework containing approximately 100 000 medically relevant items (symptoms, signs, tests, etc.). The basic principle behind the system is that the body has only a limited repertoire of reactions at its disposal. It means that one may expect predictable reactions to wide varieties of stimuli. A detailed

description of the principles of such a system has been published elsewhere [8]. A tree-based configuration, structured according to commonly used methods by which doctors interview and examine patients, permits the user to locate particular items quickly and easily. The program permits the doctor to follow his own train of thoughts by allowing him to switch from one menu to another and from one level to another. SCS actually forces the doctor to consider more closely what he knows about the patient. The patient, having a view of the screen, can immediately correct the doctor in case of misunderstanding or misinterpretation. Having located the correct item the doctor may enter this datum by pushing a single button. It guarantees unequivocally the input of a term which can certainly be recognised by the program.

Using this format the doctor is released from typewriting. 'No more typewriting' removes one of the worst impediments to computer use in routine practice and opens up the possibility of computer use simultaneously with the patient-doctor contact. It may even promote the patient's involvement in the process and foster mutual responsibility for the quality of care and cure. At any time during the process the doctor (and the patient) can check the stored data by pushing one function key and edit the data if necessary. These possibilities make paperwork rather obsolete.

The goal of diagnosis is to place a nosologic label on a process that manifests itself in the patient over time [9]. We use the word diagnosis as a precise identification of the illness of the patient as defined in the generally accepted nosology. But how the doctor arrives at such a decision is still a matter of debate. Some believe the process to be a matter of probabilistic calculation [9–12]; others take a different stance. Deber and

Baumann [13] have come to the conclusion that determining the correct diagnosis is a matter of problem solving whereas decision making refers to a situation in which a choice must be made from several possible alternatives. They support the view that problem solving in clinical medicine is based on pattern recognition following data gathering. In a previous study among 68 physicians we found the same principle [14]. Sticking as closely as possible to the doctor's reasoning process we took pattern matching as the leading strategy. This strategy looks for analogy between the patient's symptoms and the constituent elements of a predefined disease. DDSS performs this by template matching which can be executed very efficiently on a digital computer [15].

By means of the presentation of the diseases as configurations of symptoms the doctor is able to judge the value and the accuracy of his diagnosis by comparison. It will give him ample opportunity to (re)consider all possibilities offered in the differential diagnosis. This conception differs in a number of ways from most other expert or decision support systems.

1. Most (medical) knowledge systems are directed at a restriction of the possible alternatives, preferably to the best one. DDSS takes another approach. It will confront the doctor with as many explanations for the present symptom-configuration as possible. It leaves the doctor to decide which one fits best. A score number, next to every element of the system differential diagnosis, representing the degree of overlap between the patient's symptoms and the corresponding disease profile in the knowledge base, could help the doctor in his/her decision.

2. We defined 'support' in the sense of de Bono's 'lateral thinking', i.e. explore the possible alternatives and try again in various directions [16]. We have translated this idea

in the maxim: rethink—reconsider—retry—recollect—restart.

3. By the presentation of the diagnostic possibilities in their constituent composition of symptoms the DDSS can offer the doctor an encyclopedic function, which means that the finding of errors should be an ongoing concern of the examining doctor during the diagnostic process.

The knowledge base was composed of 76 disease descriptions with 60–200 items per disease. To define these descriptions close to the configurations as can be found in primary care, we constructed a blend of disease descriptions from 12–14 medical textbooks and submitted these to a panel of general practitioners. In the knowledge base each pattern is strictly defined. When definitions are used, no matter how arbitrary the definitions may seem to any particular observer, the system is guaranteed to be valid across all patients within the framework of those definitions [17]. The result was stored in the knowledge base according to the definitions of the SCS.

3. Test design

"The next step in the evaluation of computer programs will have to include examining the performance of the physician and the computer together", conclude Berner et al. [2]. This is a reasonable demand for systems for diagnostic support of which 95% did not survive the voyage from laboratory to clinic [18]. Computer-aided diagnosis has not been tested in general practice so far [7]. "Preliminary system evaluation should explore feasibility, performance, reliability and in a manner that does not put clinical subjects at risk" Miller advises [19]. We performed a small experiment in a primary care setting focusing on the following items.

1. Ease of learning. Most available computer programs are complicated and difficult to master. Costs and time expenditure for mastering a computer program sets considerable barriers for practising physicians. They ask for quick and easy handling. We concentrated on loading, use of the mouse, finding the way through the SCS and, entering and editing data.

2. Ease of operation. The first requirement is that the time with a computer does not exceed the consultation time without a computer. During the patient–doctor contact the doctor should enter and edit as many data as (s)he thinks adequate for the case at hand. Storing and checking data must run smoothly as well as the possibility to retrace steps. Diagnostic feedback can be asked for at any time during the process. Resumption of data acquisition after seeing the system's differential diagnosis must be easy.

3. Appreciation and acceptability. We collected the responses to three questions: the doctor's judgement of the computer use during consultation; the patient's acceptability of the encounter; and, the doctor's appreciation of the diagnostic decision support.

4. Methods

We invited a group of 20 doctors all linked with the Institute for Family Medicine of the Erasmus University. Requirements for candidate subjects were that they had no previous experience with the system and had not taken part in its design or implementation. The resulting group consisted of 7 women and 13 men, ages ranging from 24 to 63 years. Most subjects owned a personal computer (75%) and had some experience with some medical information system. They all took part in two teaching sessions. The first session was meant to give some background information about

the logic and the ideas behind the system. The second session was dedicated to train and experiment with the system functionally.

For the purpose of comparability we made use of simulated patients. Simulation provides a useful tool to test diagnostic systems [20]. The patient scenarios were based on real patients whose data were carefully collected, recorded and monitored throughout the illness ([8]). Eight case histories were selected to represent rather common diseases in general practice. The eight patients, of varying ages, represented eight different organ systems. Their databases contained approximately 300 items each: items ranging from social and medical background to laboratory tests and X-Ray reports. An actor played all cases convincingly (subjects' perceived realism score: 80.9 on a scale of 0–100).

Five randomly selected patients consulted the doctor within a time schedule of maximally 2 h, which means 17 min for the actual consultation, 5 min paper work (e.g. writing down diagnostic ideas, subjective probabilities, etc.) and 2 min for adaptation to the system.

5. Results

We encountered surprisingly little difficulty in teaching the system. Training for 2 h and another 2 h of actual experimenting made the participants very confident with the system and the use of the standardised terminology. In more than half the cases participants needed less time to achieve a sufficient degree of fluency in the use of the system. We clearly observed that the logical (medical) structure induced confidence in the participants that they would be able to handle it rather easily. After having done a few exercises, the task of data entry was perceived to be so easy that participants frequently lost interest. Young

participants, even if they had no computer experience at all, could be taught to use the SCS system correctly with the help of the mouse within 10 min, as we repeatedly demonstrated. The system assumes its users to be medically, not technically, trained. Twenty doctors in 99 consultations entered nearly 2000 symptoms. To find less than 20 instances where something they wish to enter could not be encoded, is a high score for the symptom coding system. The lines of free text and notepads that were available for otherwise unencodeable particulars of patients were not used.

The number of questions per consultation was 29, ranging from 25 to 33. These questions resulted in the entering of 19 symptoms. Obviously, not every question will provide a clinically relevant symptom [21]. Most questions originate from the hypothesis and the line of thinking the doctor currently entertains. In more familiar cases the questions came more easily than in more complicated ones. With these latter ones the doctor was more guided by the presented menu often stating: 'Let me see, what else can we ask'. More experienced physicians often started by asking free questions of the patient, the younger doctors were more guided by the system. In the end of each session most doctors were happy to be guided by the system which gave the data acquisition a more organised appearance. In the physical examination the physician had to use his imagination (because of the simulation mode) and, consequently (?), changed frequently from location to location and from one examination method to another (e.g. inspection, palpation). In daily practice doctors commonly use a similar mix of these examination methods. Structuring and systematisation are not the strongest points in doctors' data gathering [7,21,22].

The rather slow speed of formulating questions chiefly determined the speed of the data entering and, consequently, the consultation time. The number of questions/min hardly varied over the series: on average 1.6 question/min (ranging from 1.52 to 1.74). The actual average duration of a consultation was 16.6 min with a decrease of time expenditure after 4 patients to 13.9 min. Earlier we found an average consultation time of 11.25 min without a computer [21].

Data entering itself did not create a problem. The average speed of data entry with the system was 4–5 symptoms/min, with a maximum of 11/min. Compared with the amount of information written down by hand in a consultation (3–4 words/min) the system's possibilities exceed the conventional data recording. The overall speed increased significantly from the first to the last experimental session.

Checking and editing stored data occurred only rarely. With one keystroke a comprehensive oversight of the data stored so far could be presented. A highlighted item could be deleted or replaced with a single keystroke. This occurred in only 13 out of 1879 symptoms collected.

The number of symptoms entered, however, appeared regrettably small (17–25). This made the list of differential diagnoses rather long: 3–15 diseases with consequent low scores for matching symptoms. The correct diagnosis was present in 96% of the cases. Although the list was presented in the ranked order of matching symptoms, most participants chose the name of their working hypothesis as the diagnosis, often disregarding the score. The enthusiasm for the system's diagnostic module, however, was great. After they had completed the case and made a diagnosis on their own, they often waited with an extended finger above the pertinent function key: 'Can I press it now? Can I press

it now?', indicating a need for support and the grading of one's own performance.

If we take the precise identification of the disease as our goal then only 43% of the consultations resulted in a completely correct diagnosis being made by the doctor. In 23% of the consultations a completely incorrect diagnosis was made. In 34% of the cases the diagnoses can be seen as 'near misses'. Here we meet a stumbling-block which may be characteristic of primary care. In primary care a more generally named diagnosis, as can be found in classification systems such as ICPC, suffices to formulate a treatment plan. Vaguely defined diagnoses, however, may induce a variety of vocabularies. With Diamond et al. [17] we believe that the correct diagnosis is a very appropriate standard on which to evaluate the accuracy of any diagnostic system. After seeing the system's DD in only one case did the doctor change his diagnostic hypothesis, happily in the right direction. In only 14 cases participants started a new round of data collection after having seen the differential diagnosis. This second round, however, was usually confined to just one or two additional questions, more directed at therapy than diagnosis.

On average the participants indicated a considerable level of disturbance to the normal patient–doctor contact because of the use of the system. Interestingly, our own observations (including the 'patient's') indicate a rather poor correlation between subjective judgement and objective performance here: some participants who were very fluent in the use of the system indicated that they found its use very disturbing, while others who seemed to have much more trouble gave a more optimistic opinion. The doctors' opinions about the decision support seem to indicate that they found the system differential diagnosis rather useful, but this is in fact in contradiction with their actual behaviour.

Even in cases where a subject reached a wrong diagnosis, and when the system diagnostic support list was topped by the correct diagnosis, subjects did not change their mind about their own diagnosis [6].

6. Discussion

In constructing a medical diagnostic decision support system there are two worlds involved: medicine and informatics. It will certainly take a lot of thinking and puzzling to glue these two worlds together. Can notoriously variable and inaccurate medical data [23] be brought into harmony with the zeros and ones of the digital world? Do medical data have the potential of becoming organised and structured? Can medicine and information technology learn to understand each other? So far these efforts have not resulted in a ready-for-practice computer program that can be used in the routine practice of the (primary care) physician.

A multitude of diagnostic machines has been developed; only a few reached some maturity. Perhaps the original goal had been set too high: doctors providing the computer with the patient's data and the computer will come up with the correct diagnosis. This aspiration is still unfulfilled and will probably remain so in the foreseeable future. [1,24]. In a recent study Berner et al. [2] found that four computer programs produced the correct diagnoses in only half to three quarters of instances. These results brought Kassirer to the statement that "diagnostic computer programs have come a long way, but they still have a long, long way to go" [25].

But in the light of the figures for diagnostic accuracy of doctors these results seem less disappointing. Although Engle [24] expresses high respect regarding the "human capabilities of recognising common patterns from

observation of exceedingly large numbers of stimuli” such esteem seems rather undeserved. Diagnostic errors made by doctors have been extensively documented [26]. In the diagnostic domain de Dombal [27] found an overall accuracy for acute abdominal pain of 47.1–54.3%; Zarling et al. [28] reported an accuracy of 44% for myocardial infarction, a figure which is in the same range as in a comparable study of van der Does and Lubsen [29]; Ridderikhoff [6] found a diagnostic accuracy of 46% and the present study scored 43% for diagnostic accuracy of doctors. This sheds a different light on physicians’ capabilities.

If doctors need some assistance can computers provide the necessary help? It makes a great difference whether one defines ‘help’ as the offering of the correct answer or as a means to remind the doctor of possible errors and suggest some alternative explanations for the patient’s complaints. Most programs serve a prompting function; in this respect DDSS does not take a different viewpoint. In response to the entered data DDSS produced a list of diagnostic hypotheses like many other programs. DDSS, however, produced a percentage of 96% correct diagnoses in the test, in the sense that the correct diagnosis was presented as one of a list of diagnostic possibilities. This figure contrasts positively with the percentages of correct diagnoses in other, similar programs [2–4].

The consequence of leaving the final judgement to the doctor is to trust him/her to take the proffered advice to heart. The difference between the computer’s percentage of correct diagnoses and the physicians’ indicates a gap between the expressed need of support (‘can I press it now?’) and the actual behaviour. The computer may provide the necessary help but it is up to the doctor to act accordingly. It takes two to tango and the medical situation is not exceptional in this respect. DDSS can

furnish encyclopedic support at nearly every level and at nearly every moment during the process. It can indicate the features to look for until a diagnosis is reached (disclose a list of symptoms for each of the diagnostic names) and it can list the diagnostic possibilities related to the collected patient data so far. In the end it is up to the doctor to decide and act.

Of the three requirements: time-saving, ease-of-use and need for help, time-saving and ease-of-use are among the most precious. Nothing is more treasured by the doctor than time. Computer-aided diagnostic systems are often seen as time-consuming devices and, therefore, hardly popular in the routine of daily practice. DDSS appeared to be a program which can be used simultaneously with the patient–doctor contact. After a short practice period of just a few hours the consultation time came very close to the usual time in dealing with newly presented cases. Moreover, it is interesting to notice that not the program but the doctor was the main time determinant: the tempo of interviewing of the patient determined substantially the consultation time. The more the participant got used to the (guiding) system, the more the consultation time decreased, happily not at the expense of the data acquisition.

As was observed earlier physicians’ data acquisition appeared to be rather ineffective [21,22]. Doctors ‘extracted’ only 11% of the information as contained in the patient data bases [21]. In addition, 50% of the patient information appeared to be unrelated to the diagnostic hypotheses and remained unused during the diagnostic process. A scanty number of 19 pieces of information on average was collected per case resulting in rather long lists of differential diagnoses. As was argued earlier it was not the machine that created an impediment in entering and recording data. Improvement, therefore, must come from

better interviewing techniques by doctors and, subsequently, a more organized way of physical examination. Structured data collection has been shown to improve the diagnostic accuracy, even without a computer [7]. The system's SCS may guide the doctor through the information collection and may contribute to overall completeness. The element of structured data collection needs more attention in future tests.

Little attention has been paid to the validity and reliability of the patient data. De Dombal [22] reported that "one-sixth of patient answers were so vague that observers could not agree whether the patient had said yes or no". Using standardized terms/symptoms such problems are less likely to occur. In DDSS not the doctor's question but the proposed term is the subject of discussion. This leaves the judgement about the validity of the symptom with the only expert in the consultation—the patient. Interpretation and communication errors can thus be limited. Moreover the patient becomes involved in the process and can be held partly responsible for it. This change from subjective interpretation to objective evaluation is seen as an important step forward in the diagnostic process. The standardization brings the doctor and the patient together in using the same language; not only between doctor and patient but among doctors themselves [30]. 'If there is one single change which would assist the development of information science in medicine it would be the adoption of a standard, predefined terminology amongst doctors everywhere'.[5] Although a lot of work has been done in the field of standardization of terminology (ICD, ICPC, among others), these systems are mainly focused on standardization of diagnostic names. In our view standardization must be built bottom-up, i.e. starting with the essential features which compose the various entities as they are

classified in the medical knowledge. Just as nature basically consists of atoms and molecules, these disease-patterns form a mass of small but significant features of the patient.

Standardization of terminology would also deal with the lack of consensus among doctors about disease descriptions and, subsequently, the construction of knowledge bases which can be used around the world. Construction of knowledge bases have been compared with the building of medieval cathedrals, but by lack of consensus and standard terminology it becomes a Sisyphean task. Giuse et al. [31] conducted an extensive training program for reaching consensus among seven physicians upon one disease description. We used a less laborious method. Whether our method of building a (limited) knowledge base is valid is a matter of debate. Because of the lack of a universally accepted taxonomy of diseases no reference point can be viewed as a gold standard.

Computer-aided diagnosis has not been tested in general practice so far. This is partly due to the practical problems concerning the interface between the doctor and the computer and partly to the wide variety of often unstructured problems seen in general practice. Avoiding type-writing appeared to be of great help in learnability and acceptability of the system. For the test we have limited the 'wide variety of problems' to 8 rather common patient cases representing disease entities in 8 different organ systems. This test range may diverge from the broad scale of illnesses but gives sufficient indication of the performance of the DDSS in our opinion. But it will certainly take some time and practice for doctors to adapt to the functions and possibilities of the system and to accept DDSS as a ready-at-hand consultant.

7. Conclusion

Providing diagnostic support for primary care doctors is feasible. Because it takes place simultaneous with the consultation it may contribute to the quality of the patient–doctor encounter: defining and specifying the data and involving the patient in his/her clinical process. The availability of immediate support, the option of retracing and reconsidering particular steps, and an open discussion with the patient provide opportunities for the practising doctor to improve the quality of his/her cure and care. In such a situation the diagnostic process is clear and ready for inspection and argumentation. DDSS can make a contribution to its quality.

References

- [1] R.A. Miller, Why the standard view is standard: people, not machines, understand patients' problems, *J. Med. Philos.* 15 (1990) 581–591.
- [2] E.S. Berner, G.D. Webster, A.A. Shugerman, et al., Performance of four computer-based diagnostic systems, *N. Engl. J. Med.* 330 (1994) 1792–1796.
- [3] G.C. Sutton, How accurate is computer-aided diagnosis?, *Lancet* 2 (1989) 905–908.
- [4] A.S. Elstein, C.P. Friedman, F.M. Wolf, et al., Effects of a decision support system on the diagnostic accuracy of users: a preliminary report, *J. Am. Med. Inform. Assoc.* 3 (1996) 422–428.
- [5] F.T. de Dombal, Medical diagnosis from a clinician's point of view, *Method. Inf. Med.* 17 (1978) 28–35.
- [6] E. van Herk, A diagnostic decision support system for general practice, Thesis, Erasmus University, Rotterdam, 1994.
- [7] G.C. Sutton, Computer-aided diagnosis: a review, *Br. J. Surg.* 76 (1989) 82–85.
- [8] J. Ridderikhoff, *Methods in Medicine*, Kluwer, Dordrecht, 1989.
- [9] R.A. Miller, Medical Diagnostic Decision Support Systems—past, present, and future, *J. Am. Med. Inform. Assoc.* 1 (1994) 8–27.
- [10] R.M. Dawes, D. Faust, P.E. Meehl, Clinical versus actuarial judgement, *Science* 24 (1989) 1668–1674.
- [11] S.G. Pauker, G.A. Gorry, J.P. Kassirer, W.B. Schwartz, Towards the simulation of clinical cognition. Taking the present illness by computer, *Am. J. Med.* 60 (1976) 981–996.
- [12] W.I. Card, Mathematical method in diagnosis, *J. R. Coll. Physicians London* 9 (1975) 193–196.
- [13] R.B. Deber, A.O. Baumann, Clinical reasoning in medicine and nursing: decision making versus problem solving, *Teach Learn Med.* 4 (1992) 140–146.
- [14] J. Ridderikhoff, Medical Problem Solving: an exploration of strategies, *Med. Educ.* 25 (1991) 196–207.
- [15] W.H.E. Day, Optimal algorithms for comparing trees with labeled leaves, *J. Classif.* 2 (1985) 7–28.
- [16] E. de Bono, *New Think*, Basic Books, New York, 1967.
- [17] L.W. Diamond, V.G. Mishka, A.H. Seal, D.T. Nguyen, Are normative expert systems appropriate for diagnostic pathology?, *J. Am. Med. Inform. Assoc.* 2 (1995) 85–93.
- [18] H.P. Lundsgaarde, Evaluating medical expert systems, *Soc. Sci. Med.* 24 (1987) 805–819.
- [19] R.A. Miller, Editorial: evaluating evaluations of medical diagnostic systems, *J. Am. Med. Inform. Assoc.* 3 (1996) 429–431.
- [20] G.A. Gorry, Modelling the diagnostic process, *J. Med. Educ.* 45 (1974) 293–302.
- [21] J. Ridderikhoff, Information exchange in patient–physician encounter. A quantitative approach, *Method. Inf. Med.* 32 (1993) 73–78.
- [22] F.T. de Dombal, Computer-aided decision support in clinical medicine, *Int. J. Biomed. Comput.* 24 (1989) 9–16.
- [23] A.L. Komaroff, The Variability and inaccuracy of Medical Data, *Proc. IEEE* 67 (1979) 1196–1207.
- [24] R.L. Engle, Attempts to use computers as diagnostic aids in medical decision making: a thirty-year experience, *Perspect. Biol. Med.* 35 (1992) 207–219.
- [25] J.P. Kassirer, A report card on computer-assisted diagnosis—the grade: C, *N. Engl. J. Med.* 330 (1994) 1824–1825.
- [26] O. Nacke, G. Wagner, Bibliographie zum Thema 'Die Rolle des Fehlers in der Medizin: Fehlerforschung als Aufgabe der medizinischen Dokumentation', *Method. Inf. Med.* 3 (1964) 132–150.
- [27] F.T. de Dombal, V. Dallos, W.A.F. McAdam, Can computer aided teaching packages improve clinical care in patients with acute abdominal pain?, *BMJ* 302 (1991) 1495–1497.

- [28] E.J. Zarling, H. Sexton, P. Milnor, Failure to diagnose acute myocardial infarction, *J Am Med Assoc* 250 (1983) 177–181.
- [29] E. van der Does, J. Lubsen, Acute coronary events in general practice: the imminent myocardial infarction study, Thesis, Erasmus University, Rotterdam, 1978.
- [30] J. Ridderikhoff, The quality and reliability of medical information in the health care system, in: G. Duru, (Ed.), *La Science des Systèmes dans le Domaine de la Santé*, Masson, Lyon, 1988, pp 177–180.
- [31] N.B. Giuse, D.A. Giuse, R.A. Miller, et al., Evaluating consensus among physicians in medical knowledge base construction, *Method. Inf. Med.* 32 (1993) 137–145.