Decision support system for the diagnosis of schizophrenia disorders

D. Razzouk[1], J.J. Mari[1],
I. Shirakawa[1], J. Wainer[2]
and D. Sigulem[2]

[1]Departamento de Psiquiatria, [2]Departamento de Informática Médica,
Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo,
SP, Brasil

Abstract

Correspondence

D. Razzouk
Departamento de Psiquiatria
EPM, UNIFESP
Rua Dr. Bacelar, 334
04026-001 São Paulo, SP
Brasil
Tel/Fax: +55-11-5084-7060
E-mail: razzouk@psiquiatria.epm.br

Clinical decision support systems are useful tools for assisting physicians to diagnose complex illnesses. Schizophrenia is a complex, heterogeneous and incapacitating mental disorder that should be detected as early as possible to avoid a most serious outcome. These artificial intelligence systems might be useful in the early detection of schizophrenia disorder. The objective of the present study was to describe the development of such a clinical decision support system for the diagnosis of schizophrenia spectrum disorders (SADDESQ). The development of this system is described in four stages: knowledge acquisition, knowledge organization, the development of a computer-assisted model, and the evaluation of the system's performance. The knowledge was extracted from an expert through open interviews. These interviews aimed to explore the expert's diagnostic decision-making process for the diagnosis of schizophrenia. A graph methodology was employed to identify the elements involved in the reasoning process. Knowledge was first organized and modeled by means of algorithms and then transferred to a computational model created by the covering approach. The performance assessment involved the comparison of the diagnoses of 38 clinical vignettes between an expert and the SADDESQ. The results showed a relatively low rate of misclassification (18-34%) and a good performance by SADDESQ in the diagnosis of schizophrenia, with an accuracy of 66-82%. The accuracy was higher when schizophreniform disorder was considered as the presence of schizophrenia disorder. Although these results are preliminary, the SADDESQ has exhibited a satisfactory performance, which needs to be further evaluated within a clinical setting.

Introduction

Schizophrenia is a psychotic disorder that induces global disablement of the individual's psychosocial functioning. Although its etiology is unknown and its treatment elicits only a partial response, it is very important to have early detection of the initial symptoms, since early therapeutic intervention permits the prevention of the worst outcome. Since the clinical presentation of schizophrenia is heterogeneous, many operational diagnostic criteria have been developed during the last three decades, but there is no consensus as to which of them is the most adequate.

The International Classification of Dis-

eases, ICD-10, and the Diagnostic and Statistical Manual of Mental Disorders, DSM-IV systems, are the two diagnostic classificatory systems most used in clinical and research activities (1,2). The two systems permit us to obtain a different profile of schizophrenia. The ICD-10 permits the diagnosis of schizophrenia with only one psychotic symptom lasting one month (1). On the other hand, DSM-IV requires 6 months to diagnose schizophrenia, at least two psychotic symptoms and a psychosocial dysfunction must be present (2). The study of Bell et al. (3) illustrated this situation when they compared 11 operational diagnostic criteria for schizophrenia in a sample of 470 first-stage psychotic patients. The agreement about the presence and absence of schizophrenia between them was only 1.7 and 4.6%, respectively. These results have had considerable impact on clinical practice since they demonstrated that a subject could be diagnosed as schizophrenic according to chosen criteria and to the theoretical background of the psychiatrist.

Thus, a diagnosis made even by experts in their clinical practice is based on abstract models organized according to the physician's clinical experience, theoretical background and preference for diagnostic criteria (4,5). This can also be an important factor for the learning of diagnostic reasoning by recently graduated psychiatrists.

This problem is minimized in research practice by using polydiagnostic tools such as the Operational Criteria Checklist, OPCRIT (5-8). This system is a 90-item computerized checklist used to diagnose psychotic disorders through 12 operational diagnostic criteria (7-9). Although it is a reliable and valid instrument, it is not applicable to routine clinical practice and cannot be used to permit a student to learn how to recognize schizophrenia disorder (9,10).

Decision support systems or expert systems have been developed in medicine to assist the physician in the diagnostic deci-sion-making process (11-13). Ideally these computerized systems are designed using artificial intelligence techniques and represent the expert's reasoning in situations that require clinical problem-solving tasks. The expert's reasoning is characterized by efficient and quickly cognitive shortcuts (named "clinical reasoning skills") triggered by a few elements (14).

However, in psychiatry, such systems are rare, especially to diagnose psychotic disorders (15-17). The development of an intelligent system to diagnose schizophrenia is an important initiative because it permits the evaluation of expert clinical reasoning and its influence on the diagnostic decision-making process. Our goal is not to create or to replace the operational criteria but to create an intelligent system which uses an explicit and valid clinical model of schizophrenia.

There is some evidence that decision support systems are effective in improving learning by medical students (18,19). We suggest that expert systems could be useful to students to support their learning process because these systems show the elements employed for the diagnostic decision-making process. One of the reasons for developing an expert system is to clarify the steps of clinical reasoning for the students; however, it is not the aim of the present research to test the educational usefulness of this system.

The objective of the present study was to describe the development and evaluation of a decision support system (SADDESQ), a useful tool to help students and novice psychiatrists to understand all the necessary steps in the diagnostic decision-making process.

Material and Methods

The study comprises four principal phases: knowledge acquisition, knowledge organization, knowledge modeling, and the evaluation of the system's performance.

The first stage was the knowledge acqui-

sition process that normally involves the extraction of knowledge from one or more experts by identifying their cognitive inferences, concepts and meanings within a decision-making situation. A pilot study with three "experts" was carried out to explore the similarities of the clinical patterns used in the diagnostic decision-making process (20). Four clinical vignettes of schizophrenia were used to elicit from them how they identify schizophrenia symptoms and how they use them to diagnose schizophrenia. The graph methodology was used because it has been shown to be useful to induce experts to design graphs with the associations of symptoms they believe to be necessary to reach a diagnosis (21,22). A graph is a finite set of dots called nodes connected by links called arcs. A path is a sequence of consecutive arcs in a graph. The node was used to represent diagnoses and the arcs showed how symptoms were connected with diagnosis. Thus, the paths used by an expert to reach a diagnosis may be visualized through the graph structures. Figure 1 shows a possible graph representing the diagnosis of schizophrenia. Analysis of graph structures permits a quantitative and qualitative approach. The former measures the number of nodes and trees, the number of levels of nodes and the median values of nodes. The latter compares symptoms and their degree of specificity. Thus, the analysis of these graphs permits the identification of the triggering symptoms involved in the decision-making process. The triggering symptoms were identified as the most frequent ones and those with the highest values in the graphs. In the pilot study, the triggering symptoms were compared between the three experts (20). The process used to construct the graphs is described below.

The expert was asked to report the most important signs and symptoms of schizophrenia included in clinical case vignettes produced from the schizophrenia patients' charts containing the most complete data

from the Schizophrenia Outpatient Program of the Federal University of São Paulo. Next, the expert was asked to construct a list with the most significant symptoms identified from the vignettes. The expert then chose a group of symptoms from this list and constructed the graphs. The expert could construct as many graphs as he needed. Each graph corresponded to one possible diagnosis of schizophrenia. Then, each symptom was graded from 0 to 10 according to its specificity for the diagnosis of schizophrenia.

Since the analysis showed that there was disagreement between the three experts about the triggering symptoms, only one expert was selected as the source of knowledge (20). In the other words, the experts exhibited three different patterns of reasoning for diagnosing schizophrenia. This finding represented a hindrance in the construction of a consensual pattern of reasoning and its transposition into a coherent model. Thus, the expert who chose disorganization as the triggering symptom was chosen because this symptom was considered to be broader and simpler to describe than the others (20).

The second phase was to collect and organize data from open interviews with this expert over an 18-month period. The method for data collection was the same as that used in the first phase. Then, seven other vignettes elaborated from the schizophrenia charts were presented to the expert. A qualitative analysis was also employed to identify

Figure 1. Figure representing schizophrenia disorder, where S (S1....S3) are symptoms, signs and laboratory tests important for diagnosis (D) and N are the decision nodes. The apex node corresponds to the more specific diagnosis D2. N1 and N2 represent levels of decision, i.e., the closer to the apex, the more conclusive the decision and the higher the degree of diagnostic certainty. Each node N has a tree that represents different combinations of symptoms. Values of 0 to 10 were attributed to each symptom according to its specificity (0 = totally unspecific for diagnosis of schizophrenia disorder and 10 = very specific for diagnosis of schizophrenia disorder).
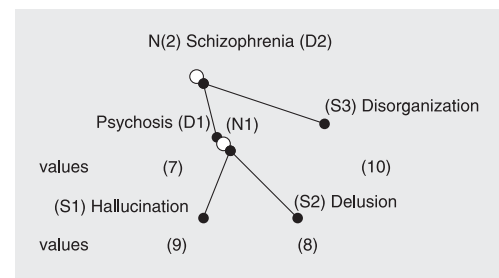
the key elements in the graphs and to construct the algorithms corresponding to each phase of the diagnostic decision-making process. For instance, we asked the expert in which conditions he could make the diagnosis of schizophrenia without the presence of disorganization, and how long a time was necessary to make a diagnosis.

A series of open interviews were also held to precisely define the concept of disorganization and to identify how the expert had reached this construct. A glossary of technical terms was also elaborated with the expert. The expert was interviewed by a psychiatrist (D. Razzouk) with expertise in the schizophrenia domain and all the interviews were recorded and transcribed in order to permit a qualitative discourse analysis. Then, the interviewer constructed operational rules to identify the disorganization symptom based on the expert's discourse. Once the concept of disorganization was represented, the algorithms for schizophrenia were constructed according to eight different clinical contexts identified in discourse analysis. Finally, all the collected and analyzed data were shown to and extensively discussed with the expert, who was asked to comment on them.

The third phase of this study was transferring the clinical model to a computational representation (knowledge modeling). The technological approach adopted in this system is based on the concept of parsimonious cover (23). The parsimonious cover theory defines a diagnosis as the smaller set of diseases that explains all symptoms known to be present. For example, if disorder D1 can cause symptoms S1, S2, and S3, and D2 can cause S3 and S5, then if a patient is known to have symptoms S2 and S3, the two plausible diagnoses are (D1) and (D2), that is to say, there are two competing diagnoses, one that states that the patient has only disease D1 and one that states that the patient has only disease D2. Both diagnoses "cover" or "explain away" all symptoms known to be present. Essentially, the system represents the potential causal connections between diseases and symptoms, and different reasoning algorithms operate on such knowledge.

The fourth and final phase was the evaluation of the system's performance (SADDESQ). The assessment of a decision support system is a complex process that involves a laboratory test and a field test (clinical setting) (24-26). The laboratory test involves the assessment of reliability and internal validity. Reliability means the same input originating the same output. Internal validity measures the agreement between the system output and the gold standard (expert). The internal validity was used to evaluate if the domain knowledge was accurately represented. Thirty-eight vignettes from the charts of the Outpatient Program of Schizophrenia and Affective Disorders were prepared according to the completeness of the data. The expert analyzed these 38 vignettes and the results were compared with the output from the SADDESQ. The expert who was the source of knowledge was considered to be the gold standard for the diagnosis of schizophrenia in 38 clinical vignettes. The expert should diagnose the following five categories as "schizophrenia present", "possible schizophrenia", "schizophreniform present", "schizophrenia absent" and "inconclusive". The categories "schizophrenia present", and "possible schizophrenia" were always considered as positive cases. The category "schizophreniform disorder" was first analyzed as a positive case and later as a negative case. Another expert (D. Razzouk) confirmed the presence of psychopathological symptoms in 38 clinical vignettes with psychotic disorder diagnoses and entered the data (answers to all nine questions) into the system (SADDESQ). This system provides output with all possible psychotic diagnoses. Cases were considered to be positive when outputs were schizophrenia as a single hypothesis or one of the possible hypotheses. Schizophreniform disorder was first consid-

ered as a positive case and later as a negative case. As the sample size was small, we analyzed only dichotomous variables (positive and negative cases of schizophrenia). The Phi coefficient (for nominal variables) was used to measure the correlation between the expert and SADDESQ and the Cohen kappa coefficient was used to measure their agreement.

Results

The first phase concerning the disagreements among the three experts was described elsewhere (20). The second phase consisted of exploring the concepts of the expert about disorganization and to organize the data collected. The expert constructed 19 graphs representing the diagnosis of schizophrenia. Disorganization was the crucial symptom for his decision process; however, this concept was used in a broader way than found in the psychiatric literature. This expert defines disorganization as a group of four symptoms: negative symptoms, social and interpersonal dysfunction, permanent and progressive changes in patients' personality, and inadequate behavior. During the discourse analysis, it was clear that disorganization was a longitudinal concept, i.e., it was recognized as the result of a transformation process in the subject's life, personality and social rapport. Thus, the expert was able to recognize disorganization in the first psychotic episode only if the symptoms had been present for more than 1 month but less than 6 months. Some rules were developed to identify the presence of the four groups of symptoms. Then, it was agreed that if two of these groups of symptoms were present, disorganization would be considered to be present. The second step was to organize all data collected through algorithms considering eight clinical contexts: with or without drug abuse; with or without disorganization symptoms; with one or multiple psychotic episodes; with or without affective symptoms; with or without organic causes; with or without socio-occupational impairment; with or without recurrence of psychotic symptoms, and the duration of illness. These situations were considered on the basis of the algorithms because they are present in all operational diagnostic criteria to distinguish schizophrenia from other mental illnesses.

The SADDESQ software comprises nine questions (to be answered: yes or no, or I don't know) about mood disturbances (presence/absence and predominance of mood/ psychoses), duration of psychotic symptoms, drug use pattern, number of psychotic episodes, socio-occupational dysfunction, presence of identifiable organic causes, and disorganization. SADDESQ differentiates eight diagnoses of psychotic disorders: schizophrenia, schizoaffective, schizophreniform, brief psychotic disorder, mood disorders, psychosis due to drug use, delusional disorder, and other psychotic disorders. The system has a graphic interface that allows the user to select which question to answer next, but using the reasoning concerning the usefulness of each unanswered question, it indicates which unanswered question is crucial for the final diagnosis. If the competing diagnoses include diseases of different "severity", the system informs the user that it is not yet possible to make a firm diagnosis and which are the missing data that could eliminate one of the hypotheses.

The fourth stage included the evaluation of the internal validity of the SADDESQ. The results of the expert and those obtained by the SADDESQ exhibited a moderate to good correlation and level of agreement (r = 0.39-0.64; kappa = 0.35-0.63). The data in Tables 1 and 2 indicate that the SADDESQ correctly diagnosed the presence of schizophrenia in 85-89% of the true cases. Nevertheless, the SADDESQ incorrectly diagnosed the absence of schizophrenia in 26-44% of the negative cases. The misclassification rate was higher when schizophreniform disorder was considered as the absence of schizophrenia disorder. Schizophreniform disorder is a provisory di-

agnosis made when the psychiatrist does not have sufficient information to decide about schizophrenia disorder in the first psychotic episode. There were six cases in which the expert diagnosed the schizophreniform disorder and the SADDESQ diagnosed it as the schizophrenia disorder. In summary, SADDESQ exhibited a low rate of misclassification (18-34%), with an acceptable accuracy (66-82%; Tables 1 and 2).

Discussion

The methodology needed to develop an expert system is complex and controversial.

Table 1. Comparison of the diagnoses resulting from analysis of 38 vignettes between the SADDESQ and the expert when schizophreniform disorder was considered to correspond to the presence of schizophrenia.

	Expert (gold standard)		Total
	Schizophrenia present	Schizophrenia absent	
SADDESQ			
Schizophrenia present	17 (89%)	5 (26%)	22 (58%)
Schizophrenia absent	2 (11%)	14 (74%)	16 (42%)
Total	19 (100%)	19 (100%)	38 (100%)

SADDESQ = clinical decision support system for the diagnosis of schizophrenia spectrum disorders. Accuracy = 0.81; misclassification rate = 0.18; correlation coefficient (phi) = 0.64 (P = 0.01); kappa Coehen's coefficient = 0.63 (asymptotic standard error = 0.12, P < 0.000, chi-square test).

Table 2. Comparison of the diagnoses resulting from the analysis of 38 vignettes between the SADDESQ and the expert when schizophreniform disorder was considered to correspond to the absence of schizophrenia.

	Expert (gold standard)		Total
	Schizophrenia present	Schizophrenia absent	
SADDESQ			
Schizophrenia present	11 (85%)	11 (44%)	22 (58%)
Schizophrenia absent	2 (15%)	14 (56%)	16 (42%)
Total	13 (100%)	25 (100%)	38 (100%)

SADDESQ = clinical decision support system for the diagnosis of schizophrenia spectrum disorders. Accuracy = 25/38 = 0.66; misclassification rate = 0.34; correlation coefficient (phi) = 0.39 (P = 0.01, chi-square test); Cohen's kappa coefficient = 0.35 (asymptotic standard error = 0.13, P < 0.01, chi-square test).

The usefulness and the impact of such systems in medicine are unclear (11,27). In Psychiatry there are additional hindrances such as the lack of valid constructs of mental disorders, the subjective assessment of psychiatric symptoms, schizophrenia as a heterogeneous phenomenon, the absence of biological markers, and finally the absence of a gold standard (5,28-30).

We would like to point out two main issues: the consequences of developing a model based on one expert and the evaluation of the system. The first issue deals with how to select the best knowledge available concerning the diagnosis of schizophrenia disorder in order to construct a knowledge base. The operational criteria and structured interviews do not allow psychiatrists to recognize schizophrenia but help them to classify it (29,30). Psychiatrists differ in their theoretical background and psychopathology concepts. Although clinical reasoning is imperfect, it is the only instrument psychiatrists use to diagnose their patients in clinical practice. The development of an intelligent system based on clinical reasoning may be questionable but the most important contribution of such a tool is the possibility to explore and test the validity of the expert's beliefs concerning schizophrenia. The construction of SADDESQ has permitted to formalize, to organize and to test informal knowledge based on clinical expertise. Considering a chaotic scenario with multiple definitions of schizophrenia and without a clue about which one is correct, this effort represents a valuable initiative to evaluate the validity of one source of knowledge. We should emphasize that our students and our patients are exposed to imprecise and sometimes invalid information. Thus, it is important to understand how multiple concepts of schizophrenia are integrated into clinical reasoning. The students develop their clinical reasoning mostly with the experts and this acquired knowledge may or may not be valid; therefore this kind of knowledge is not tested

or evaluated (31). SADDESQ contains one pattern of clinical "reasoning" about schizophrenia that can be evaluated against other sources of knowledge.

The limitations of the present study are mostly related to low generalizeability, since a qualitative approach was used to collect the data. Nevertheless, generalizeability was not the goal of this study, because the crucial point was to identify the patterns of reasoning used by different experts and to determine if the experts share common reasoning shortcuts in this decision-making situation. Since they did not show a clear common path of decision it was preferable to concentrate on one expert and to explore deeply one or more patterns of reasoning. Experts may reach the same diagnosis by different pathways. Then, the question was to know which of these reasoning shortcuts were more feasible to construct a clinical model of schizophrenia.

The methodology for knowledge acquisition is still an open question (12,32,33). There are two theoretical approaches to the knowledge acquisition domain: the first derived from Cognitive Sciences and the second from Mathematical and Logical Models. Cognitive Science involves the study of concepts and the analysis of protocol languages and of the psychological aspects of discourses. However, available complex techniques derived from mathematical models do not solve problems such as the usefulness and usability of such systems. In other words, if the end-user does not understand how the data output was processed in terms of neural network systems, the system will probably be abandoned. These systems are useful in situations in which humans have difficulty to calculate or to process a large number of variables. However, in the diagnostic decision context, the problem is not to process quantity but to identify the quality and relevance of information within a changeable context. In the first and second phases of this study we employed the cognitive approach but in the third phase a mathematical approach was used for modeling knowledge. Our concerns were to create a friendly and simple clinical tool to help the end-user, which, however, should be tested in future studies.

The consequences of selecting one expert to construct a model are associated with his theoretical background and expertise, i.e., this permits increasing or diminishing the number of diagnostic hypotheses (diagnostic bias). Although the experts exhibited different patterns of reasoning, this does not mean that the resulting model represented by SADDESQ does not share concepts with other experts from different schools. However, our choice was based on the pattern that would provide a broad concept of schizophrenia.

The knowledge acquisition phase is strongly influenced by the state of the art of knowledge concerning schizophrenia. Moreover, it is important to emphasize that the process of acquiring knowledge was facilitated because the expert was interviewed by another psychiatrist (D. Razzouk) also with expertise in the schizophrenia domain.

The qualitative discourse analysis showed that the expert's concept about schizophrenia was mainly influenced by Kraepelin's and Bleuler's theoretical concepts because the concept of disorganization involves negative symptoms and signs of dysfunctional behaviors. These theoretical concepts are based on the poor prognosis outcome and the identification of cognitive deficits. We must also emphasize that the concept of disorganization developed herein seems to be closely linked with psychopathological dimensional models described in the literature (5). Although the beginning of these symptoms is slow and progressive, the expert could recognize schizophrenia even before the symptoms became evident. Thus, the identification of the elements triggering his reasoning was useful for the construction of a model that permits the diagnosis of

schizophrenia even without all of the symptoms required by traditional operational criteria. For instance, the DSM-IV system adopts a narrow concept of schizophrenia that requires at least 6 months of symptoms and the completion of the first psychotic episode and socio-occupational dysfunction (2). In contrast, the reasoning of psychiatrists tends to be more comprehensive and adaptable to the clinical heterogeneity of schizophrenia. SADDESQ allows diagnosing schizophrenia, even in the first psychotic episode, before completing 6 months of psychotic symptoms. This flexibility in the diagnostic process is closer to what occurs in clinical practice. Thus, we suggest that this decision support system for the diagnosis of schizophrenia is likely to be less narrow than the DSM-IV system. However, this will be determined in a future study by comparing the SADDESQ with a structured instrument such as the OPCRIT or a panel of experts.

The other issue is how to evaluate the expert system (24-26,34). Wyatt (26) suggested three fundamental measures to evaluate these systems: structure, performance and impact. The assessment of the structure means that the system should contain the correct knowledge. The performance means that the system runs adequately (speedy, accurate, etc.). Usually, for this phase, vignettes can be used in order to compare the results of the diagnoses generated by the system and those of the gold standard (expert). The impact (external validity) of the system must be measured so as to assess its efficacy and effectiveness regarding the physician's decisions and consequently patient care. In the present study, we describe only the evaluation of the performance of the system.

The absence of a gold standard to establish diagnostic validity has been the focus of the problem of how to test these tools (24-26,34). We would argue that the type of diagnostic criterion adopted as a gold standard influences the measurement of the validity of the expert systems. Because we must know if the computerized model agrees with the expert-based model (i.e., if domain knowledge is accurately represented), the gold standard selected was the expert who was the principal source of the knowledge acquisition phase. The results regarding the performance of SADDESQ are promising because the misclassification rate was low (Tables 1 and 2). However, it is important to point out the tendency of SADDESQ to identify schizophrenia, when absent, resulting in a lower specificity. Part of this misclassification is not an unacceptable error because the schizophreniform disorder is a provisory diagnosis that can frequently be changed to the diagnosis of schizophrenia one year later (Tables 1 and 2). These validity parameters are still preliminary and they are not sufficient to assess the global performance of the system (Tables 1 and 2). Therefore, vignettes were used to evaluate these systems leading to additional bias, lower reliability between psychiatrists to recognize psychopathological symptoms and insufficient information to allow an expert to make a valid diagnosis. The sample size of vignettes was also too small to detect more detailed differences concerning the eight categories of psychotic disorders. The complete evaluation of this system should be made in a real clinical context (to avoid diagnosis and interpretation bias), against external validations (expert panels or other intelligent systems as gold standard) and obviously, by selecting a more representative sample. This system has good internal consistency on the basis of comparison with an expert and thus it is now ready to be evaluated (external validity) more rigorously in a real clinical context.

Conclusion

The limitations of this study mainly concern the knowledge acquisition phase in which a qualitative approach was used to

collect data based on a single expert. As a result, the data collected and the final model of schizophrenia have a low generalizeability.

Further evaluations comparing SADDESQ with other sources of knowledge (structured interviews or other expert systems) will be able to measure rigorously its validity and generalizeability in the clinical context as well to evaluate its useability as educational tool.

Acknowledgments

References

1. World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines*. World Health Organization, Geneve, Switzerland.

2. American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. American Psychiatric Association, Washington, DC, USA.

3. Bell RC, Dudgeon P, McGorry PD et al. (1998). The dimensionality of schizophrenia concepts in first episode psychosis. *Acta Psychiatrica Scandinavica*, 97: 334-342.

4. Edlund MJ (1986). Causal models in psychiatric research. *British Journal of Psychiatry*, 148: 713-717.

5. Peralta V & Cuesta MJ (2000). Clinical models of schizophrenia: a critical approach to competing conceptions. *Psychopathology*, 33: 252-258.

6. Berner P, Katschnig H & Lenz G (1986). First-rank symptoms and Bleuler's basic symptoms. New results in applying the polydiagnostic approach. *Psychopathology,* 19: 244-252.

7. McGuffin P, Farmer A & Harvey I (1991). A polydiagnostic application of operational criteria in studies of psychotic illness: development and reliability of the OPCRIT system. *Archives of General Psychiatry*, 48: 764-770.

8. Dolfus S & Brazo P (1997). Clinical heterogeneity of schizophrenia. *Psychopathology*, 30: 272-281.

9. Azevedo MH, Soares MJ, Coelho I et al. (1999). Using consensus OPCRIT diagnoses: an efficient procedure for best-estimate lifetime diagnoses. *British Journal of Psychiatry*, 175: 154-157.

10. Forrester A, Owen DGC & Johnstone EC (2001). Diagnostic stability in subjects with multiple admissions for psychotic illness. *Psychological Medicine*, 31: 151-158.

11. Miller R (1994). Medical diagnostic decision support systems - past, present and future: a threaded bibliography and brief commentary. *Journal of the American Medical Informatics Association*, 1: 8-27.

12. Patel VL & Kaufman DR (1998). Science and practice: a case for medical informatics as a local science design. *Journal of the American Medical Informatics Association*, 5: 489-492.

13. Szolovits P, Patil RS & Schwartz WB (1988). Artificial intelligence in medical diagnosis. *Annals of Internal Medicine*, 108: 80-87.

14. Norman GR (2000). The epistemology of clinical reasoning. *Academic Medicine*, 75 (Suppl): S127-S135.

15. Ohayon M (1987). Intelligence artificielle et psychiatrie: modélisaiton du raisonnement diagnostique. *Annales Medico-Psychologiques*, 145: 483-502.

16. Amaral MB, Satomura Y, Honda M et al. (1995). A psychiatric diagnostic system integrating probabilistic and categorical reasoning. *Methods of Information in Medicine*, 34: 232-243.

17. Razzouk D, Shirakawa I & Mari JJ (2000). Sistemas inteligentes no diagnóstico da esquizofrenia. *Revista Brasileira de Psiquiatria*, 22 (Suppl I): 39-41.

18. Leung GM, Johnston JM, Tin KY et al. (2003). Randomised controlled trial of clinical decision support tools to improve learning of evidence based medicine in medical students. *British Medical Journal*, 327: 1090.

19. Friedman CP, Elstein AS, Wolf FM et al. (1999). Enhancement of clinicians' diagnostic reasoning by computer-based consultation. *Journal of the American Medical Association*, 282: 1851-1856.

20. Razzouk D, Mari JJ, Shirakawa I et al. (2003). Knowledge acquisition in schizophrenia: clinical reasoning patterns among three experts. *Schizophrenia Research*, 63: 295-296.

21. Leão BF & Rocha AF (1990). Proposed methodology for knowledge acquisition: a study on congenital heart disease diagnosis. *Methods of Information in Medicine*, 29: 30-40.

22. Bondy JA & Murty USR (1982). *Graph theory with applications* [http://www.ecp6.jussieu.fr/pageperso/bondy/bondy.html]. Accessed April 1, 2005.

23. Mitchell T (1997). *Machine Learning.* McGraw-Hill, New York.

24. Friedman CP & Wyatt J (1997). Development measurement technique. In: Friedman CP & Wyatt J (Editors), *Evaluation Methods in Medical Informatics*. Springer-Verlag, New York, 119-153.

25. Trochim W (1999). *The Research Methods Knowledge Base*. 1st edn. Atomic Dog Publishing, Cincinnati, OH, USA.

26. Wyatt J (1997). Quantitative evaluation of clinical software exemplified by decision support systems. *International Journal of Medical Informatics*, 47: 165-173.

27. Delaney BC, Fitzmaurice DA, Riaz A et al. (1999). Can computerized decision support systems deliver improved quality in primary care? Interview by Abi Berger. *British Medical Journal*, 319: 1281.

28. Berrios GE & Chen YH (1993). Recognizing psychiatric symptoms: relevance to the diagnostic process. *British Journal of Psychiatry*, 163: 308-314.

29. Stromgren E (1992). The concept of schizophrenia: the conflict between nosological and symptomatological aspects. *Journal of Psychiatric Research*, 4: 237-246.

30. Tsuang MT, Stone WS & Faraone SV (2000). Toward reformulating the diagnosis of schizophrenia. *American Journal of Psychiatry*, 157: 1041-1050.

31. Elstein AS & Schwarz A (2002). Clinical problem solving and diagnostic decision-making: selective review of the cognitive literature. *British Journal of Medicine*, 324: 729-732.

32. Forsythe DE & Buchanan BG (1993). Knowledge acquisition for expert systems: some pitfalls and suggestions. In: Buchanan BG & Wilkins DC (Editors), *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Experts Systems*. Morgan Kaufmann Publishers, Inc., San Mateo, CA, USA, 117-124.

33. Shaw ML & Woodward JB (1993). Modeling expert knowledge. In: Buchanan BG & Wilkins DC (Editors), *Readings in Knowledge Ac-* *quisition and Learning: Automating the Construction and Improvement of Experts Systems*. Morgan Kauffmann Publishers, Inc., San Mateo, CA, USA, 78-91.

34. Smith AE, Nugent CD & McClean SI (2003). Evaluation of inherent performance of intelligent medical decision support systems: utilizing neural networks as an example*. Artificial Intelligence in Medicine*, 27: 1-27.