



Effectiveness of the Quick Medical Reference as a diagnostic tool

Jane B. Lemaire, MD; Jeffrey P. Schaefer, MD;
Lee Ann Martin, MD; Peter Faris, MSc;
Martha D. Ainslie, MD; Russell D. Hull, MB, MSc

Abstract

A NUMBER OF COMPUTER-BASED SYSTEMS with diagnostic capabilities have been developed for internal medicine. Quick Medical Reference (QMR) is one such program. The authors describe key features of QMR and report on their study of its effectiveness as a diagnostic tool. They investigated how frequently the correct diagnosis would appear among the 5 highest ranked diagnoses generated by QMR. The charts of 1144 consecutive patients admitted to a teaching unit were retrospectively screened. Eligible cases included those referred for investigation of an undiagnosed illness with an objectively proven final diagnosis ($n = 154$). Two physicians familiar with, but not experts in, the use of QMR entered clinical information abstracted from the patients' charts into the program. Physician A obtained the correct diagnosis in 62 (40%) of the 154 cases, and physician B was successful in 56 (36%) of the cases. The authors use study cases to illustrate QMR's strengths and weaknesses.

Within the practice of internal medicine the diagnosis of disease remains the cornerstone of therapeutics and management. Can computer technology mimic the clinician's integration of clinical information to formulate a differential diagnosis? A number of computer-based systems with diagnostic capabilities have been developed for internal medicine. Examples include Dxplain, Iliad, Meditel and Quick Medical Reference (QMR). These programs vary with respect to the size of their clinical databases, but all can process entered data into lists of diagnostic possibilities using algorithms. For example, Iliad and Meditel are based on Bayes' theorem; that is, they consider the pretest probability of disease and, by assigning test characteristics (sensitivity and specificity), a post-test probability of disease is estimated. Dxplain and QMR use non-Bayesian algorithms that focus on the associations between case findings (signs, symptoms, laboratory data) and individual diseases to derive a weighted assessment of a patient's clinical presentation.

Diagnostic programs may achieve a level of sophistication that influences physician behaviour. Therefore, evaluation of this technology is important.¹⁻⁶ In this article we review the QMR and present our findings of a study evaluating its effectiveness as a clinical diagnostic tool. We also examine how the software functions when challenged with cases of varying clinical complexity.

What can QMR do?

QMR is a multifaceted computer program for internal medicine.⁷⁻¹¹ It can generate a differential diagnosis from clinical information entered into its program, offer information on over 600 diseases, describe associated disorders and complications of diseases, offer strategies to confirm or exclude disorders, and provide simulated cases for educational purposes.

Differential diagnoses

When provided with clinical information, QMR will generate a list of potential diagnoses. Each is given a weighted score based on the number of diseases in the set

Education

Éducation

From the Departments of
Medicine and Community
Health Sciences, University of
Calgary, Calgary, Alta.

This article has been peer reviewed.

CMAJ 1999;161(6):725-8



(the more diseases, the lower the score), the strength of the association between individual diseases or between individual diseases and case findings, and the relative prevalence of the disease. The diagnosis with the highest weighted score is the most likely one. In general, diseases with scores under 60 are less likely to represent an accurate diagnosis.

For example, when seeking the differential diagnosis in a case, a clinician might enter the following information: female, age 30 years, arthralgia, fever, chest pain exacerbated with breathing, malar rash, weight loss, splenomegaly, thrombocytopenia, moderately elevated creatinine level, gross hematuria and a positive antinuclear antibody reaction. The program generates a list of 38 possible diagnoses, with the 5 most likely ones and their weighted scores in parentheses as follows: systemic lupus erythematosus (214), hemolytic uremic syndrome (208), acute interstitial nephritis (189), non-Hodgkin's lymphoma (185) and Goodpasture's syndrome (184).

Critiques

QMR can critique any generated diagnosis by highlighting all the findings consistent or inconsistent with the diagnosis. It can offer suggestions for further questions to confirm or exclude the diagnosis, or it can suggest potentially interesting, related diagnoses.

In the example described in the preceding section, QMR suggests that all but the findings of gross hematuria and moderately increased creatinine level are consistent with the diagnosis of systemic lupus erythematosus. It also suggests that laboratory testing for anti-double-stranded DNA, complement levels and anticardiolipin antibody may help to confirm the diagnosis. Lupus nephritis is suggested as a related diagnosis.

Disease information

QMR offers information on over 600 diseases. Many tools are available to use the data, including a disease profile database in which QMR offers disease-specific case findings, a differential diagnosis index, a disease-association database and a library of suggestions to confirm or exclude diagnoses.

Again, in the example mentioned earlier, the disease profile for systemic lupus erythematosus includes classic history, and physical and laboratory data. The QMR differential diagnosis includes rheumatoid arthritis, Lyme disease, polymyalgia rheumatica and progressive systemic sclerosis, and the associated disorders given are lupus nephritis and cerebritis.

Evaluation of QMR's effectiveness

We studied the effectiveness of QMR (version 2.2.2 [1994 update]; University of Pittsburgh, Camdat Corporation and The Hearst Corporation; 1993) as a diagnostic tool. Our primary objective was to determine how often

the correct diagnosis would appear among the 5 highest ranked diagnoses generated by QMR when the program was provided with clinical and laboratory information available to a general internist at the time of referral for investigation of an undiagnosed illness.

Methods

The study was carried out during 1995/96 and was approved by the Conjoint Research Ethics Board of the University of Calgary. The study population included patients admitted to a medical teaching unit in a tertiary care hospital (Foothills Hospital) from 1991–1993. Cases were eligible if the patient was admitted for the investigation of an undiagnosed illness that began within 6 months of presentation and for which a definitive diagnosis was confirmed.

Two of us (J.B.L. and J.P.S.) abstracted clinical information from the patients' charts. Clinical impressions noted by the admitting physicians were censored. Two others (L.A.M. and M.D.A.) who had no involvement with the program's inception or development were designated as non-expert users (physicians A and B). They reviewed the cases and entered findings into the QMR program. They worked independently and were blinded to the final diagnosis. The differential diagnoses generated by the QMR (in order of diagnostic likelihood) was compared with the final diagnosis.

Results

A total of 154 eligible cases were identified from 1144 consecutive cases reviewed. The mean age was 57.7 years (range 14–92 years), and 79 were males. Table 1 summarizes the diagnoses by disease category and their presence in the software knowledge base (QMR-KB). The correct diagnosis was present in the QMR-KB in 137 (89%) of the 154 cases. A number of diagnoses occurred more than once; for example, there were 12 cases of pulmonary thromboembolism. Of the 73 individual diagnoses observed in this study, 57 (78%) were present in the QMR-KB.

Using QMR physician A obtained the correct diagnosis in 62 (40%) of the 154 cases (95% confidence interval [CI] 32%–48%), and physician B was successful in 56 (36%) of the cases (95% CI 29%–45%); the difference was not significant ($p = 0.43$). When only the 137 cases whose correct diagnosis was in the QMR-KB were considered, the rate of diagnostic success increased to 45% (95% CI 37%–54%) for physician A and 41% (95% CI 33%–50%) for physician B.

The following are samples of the cases used in the study that illustrate the program's function.

Rare disease: A 75-year-old man presented with a fluctuant chest-wall mass, cough, fever, chest pain, dyspnea and hoarseness. He had leukocytosis and anemia, and a chest radiograph showed a pleural mass, nodules and hilar densities. QMR ranked the correct diagnosis of thoracic actinomycosis as the highest in the differential diagnosis, with a weighted score of 303. The other possible diagnoses (listed



in descending order of likelihood) were pulmonary lymphoma, Wegener's granulomatosis, squamous cell cancer of the lung, lymphomatoid granulomatosis, pulmonary nocardiosis and a pleural malignant mesothelioma.

Nonspecific presentation: A 42-year-old woman presented with a 10-day history of nausea, myalgia, fever and chills, headache, pain in the lower abdominal region and anorexia. She had hepatomegaly, leukocytosis, mild anemia and a tender abdomen. The QMR generated a differential list of over 200 diagnoses, with the correct diagnosis, tubo-ovarian abscess, listed as number 156.

Classic presentation: In the case of systemic lupus erythematosus described earlier in the article, the woman had a classic presentation and the QMR agreed. An experienced internist would probably not seek help with the diagnosis in such a case, whereas a medical student might do so.

Disease not in QMR knowledge base: A 45-year-old woman

with a remote history of systemic lupus erythematosus presented with a thickened erythematous rash over both shins. Biopsy confirmed lupus panniculitis. At the time of our study this diagnosis was not in the QMR knowledge base.

Case with comorbidity: A 54-year-old man with multiple complications of diabetes mellitus presented with a 3-day history of arthritis, fever, night sweats and progressive obtundation. The QMR generated a list of 85 possible diagnoses, none of which was the correct one (*Staphylococcus aureus* bacteremia and septic arthritis). In this case, the computer seemed confused by the multiple but nonspecific presenting complaints, whereas in reality a clinician would immediately recognize the signs of systemic infection.

Interpretation

Is QMR an effective diagnostic tool? QMR suggested

Table 1: Summary of cases used to evaluate the Quick Medical Reference (QMR) as a diagnostic tool, by disease category

Disease category	Total no. of cases <i>n</i> = 154	No. (and %) of cases in QMR knowledge base <i>n</i> = 137	Diagnoses*
Pulmonary	33	33 (100)	Pulmonary embolism (12), bacterial pneumonia (10), empyema (3), <i>Pneumocystis carinii</i> pneumonia (3), non-small-cell bronchogenic carcinoma (2), small-cell carcinoma (2), pulmonary abscess
Renal	19	19 (100)	Gram-negative urosepsis and bacteremia (10), pyelonephritis (2), renal cell carcinoma (2), diffuse proliferative glomerulonephritis, membranous glomerulonephritis, renal abscess, renal artery stenosis, ureteric stone
Systemic	18	13 (72)	Bacteremia (3), sarcoidosis (3), systemic lupus erythematosus (2), chronic salicylate toxicity (2), cytomegalovirus, lupus panniculitis , malaria, Parvovirus , polymyositis, soft-tissue abscess , tuberculous polyserositis, Wegener's granulomatosis
Cardiovascular	17	16 (94)	Bacterial endocarditis (5), acute ischemic heart disease (3), congestive heart failure secondary to dilated cardiomyopathy (3), malignant pericardial effusion (2), pericarditis (2), complete heart block , mediastinal abscess
Vascular	13	11 (85)	Deep vein thrombosis (11), arterial thromboembolism to limb, extrinsic compression of iliac vein
Hematologic	12	12 (100)	Secondary anemia (5), lymphoma (2), myeloma (2), immune thrombocytopenia (2), chronic lymphocytic leukemia
Gastrointestinal	9	6 (67)	Adenocarcinoma of colon (2), esophagitis, <i>Giardia lamblia</i> diarrhea, pancreatic abscess , pancreatitis, squamous cell carcinoma of upper gastrointestinal tract, tracheoesophageal fistula , vasculitis of gastrointestinal tract
Neurologic	9	9 (100)	Ischemic stroke (3), hemorrhagic stroke (2), bacterial meningitis, Creutzfeldt-Jakob disease, secondary neoplasm of cerebrum, toxoplasmosis of cerebrum
Hepatic	6	5 (83)	Adenocarcinoma with metastases to liver, ascending cholangitis, hepatic artery aneurysm , hepatic abscess, hepatitis B, papillary stenosis
Joint	6	5 (83)	Septic arthritis (5), septic pretibial bursitis
Endocrine	4	4 (100)	Addisonian crisis (2), Grave's disease (2)
Gynecologic	2	1 (50)	Pelvic abscess (space of Retzius) , tubo-ovarian abscess
Infectious lymphadenopathy	2	2 (100)	Infectious mononucleosis (2)
Metabolic	2	1 (50)	Diabetic ketoacidosis, hypomagnesemia
Male reproductive	1	0	Testicular abscess
Dermatologic	1	0	Bullous pemphigoid

*Diagnoses in bold represent those not present in the QMR knowledge base at the time of the study.



the correct diagnosis as one of the 5 highest ranked diagnoses in 36%–40% of the cases. We felt that QMR had limited effectiveness as a diagnostic tool in this setting. After excluding cases absent from the knowledge base, QMR's success rate improved to 41%–45%. Expansion of the knowledge base will improve the program's diagnostic success rate.¹² Our success rate was comparable to that in a study by Berner and colleagues,¹³ in which the correct diagnosis was among the 5 highest ranked diagnoses generated by QMR in 35% of cases.

The correct diagnosis was ranked highest in 23% of the cases in our study and was listed at any rank in 65%. Over 150 diagnoses were suggested in some cases. We felt that a differential of 5 diagnoses is a practical number for a clinician to investigate. A review of 150 diagnoses may lead to new diagnostic impressions, but it may also lead to excessive investigations.

Expert systems are designed to help clinicians with diagnostic problems. Cases that may not be diagnostically challenging to some physicians were not excluded from our study. QMR's diagnostic success rate was low for rare diseases, primarily because of their absence from the knowledge base. A number of study cases had comorbidities. This created a special challenge for the program, and inclusion of these cases may have negatively biased the diagnostic success of QMR.

Common diseases were well represented in our case series. This may have had a favourable impact on the diagnostic success of QMR in our study for a number of reasons, including well-developed diagnostic algorithms and improved ability of clinicians to document common conditions.

Another bias in our study may have arisen from the quality and amount of diagnostic information entered into the program. The diagnostic success will be improved when the specificity of the diagnostic information is high. In a study by Caratozzolo,¹⁴ the correct diagnosis was among the 5 highest ranked diagnoses generated by QMR in 93% of 27 clinical pathological cases published in the *New England Journal of Medicine* when all the final information about the case, including history, findings on physical examination and results of routine laboratory tests and investigations were entered into the computer. In our study the clinical information used was limited to that available to the consultant at the time of the initial referral, which would explain why our rate of diagnostic success was lower. Clinicians require assistance with a differential diagnosis early on and are not likely to enlist computer aid once the diagnosis is clear.

QMR is capable of further interactions with the user. The clinician may ask the program to offer criteria used to include or exclude a diagnosis, to present profiles about the diagnostic possibilities or to suggest appropriate investigations. Further studies are needed to determine whether the use of QMR in an iterative fashion will improve its effectiveness as a clinical tool. QMR's educational value has also been studied.^{10,11} In a recent study of QMR as a teaching

tool, Ikechukwu and colleagues¹⁵ found that although the diagnostic accuracy of interns and chief residents was greater than that of QMR, the program increased their understanding of disease processes and offered educational value.

Conclusion

We felt that QMR had limited effectiveness as a diagnostic tool in our study setting. Computer technology is now an integral part of the practice of medicine. It is important to evaluate critically the role of this technology in patient care. Expert diagnostic systems are rapidly evolving, and re-assessment of these technologies will provide exciting prospects for future research.

We thank Drs. Chris Brown, Stephen Edworthy, Rollin Brant and Laura McDougall for many helpful discussions. We also thank Gwen Schaefer and Joyce Lum for their contributions to data management and the Health Records Department at Foothills Hospital.

This study was supported by a grant from the Foothills Hospital Centre for Advancement of Health.

Competing interests: None declared.

References

1. Hilden J, Habbema JDF. Evaluation of clinical decision aids — more to think about. *Med Inf (Lond)* 1990;15:275-84.
2. Miller PL, Sittig DF. The evaluation of clinical decision support systems: what is necessary versus what is interesting. *Med Inf (Lond)* 1990;15:185-90.
3. Rossi-Mori A, Pisanelli DM, Ricci FL. Evaluation stages and design steps for knowledge based systems in medicine. *Med Inf (Lond)* 1990;15:191-204.
4. Wyatt J, Spiegelhalter D. Evaluating medical expert systems: What to test and how? *Med Inf (Lond)* 1990;15:205-17.
5. Sumner II W, Shultz EK. Expert systems and expert behavior. *J Med Syst* 1992;16:183-93.
6. Wolfram DA. An appraisal of Internist-I. *Artif Intell Med* 1995;7:93-116.
7. Miller RA, Pople HE, Myers JD. Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982;307:468-76.
8. Miller RA, McNeil MA, Challinor SM, Masarie FE Jr, Myers JD. The Internist-I/Quick Medical Reference project-status report. *West J Med* 1986;145:816-22.
9. Bankowitz RA, McNeil MA, Challinor SM, Parker RC, Kapoor WN, Miller RA. A computer-assisted medical diagnostic consultation service. *Ann Intern Med* 1989;110:824-32.
10. Bacchus CM, Quinton C, O'Rourke K, Detsky AS. A randomized crossover trial of Quick Medical Reference (QMR) as a teaching tool for medical interns. *J Gen Intern Med* 1994;9:616-21.
11. Bacchus CM, Morgan M, O'Rourke K, Detsky AS. A randomized trial to assess Quick Medical Reference as a teaching tool for medical students [abstract]. *J Clin Invest Med* 1996;19:S51.
12. Giuse DA, Guise NB, Miller RA. Evaluation of long-term maintenance of a large medical knowledge base. *J Am Info Assoc* 1995;2:297-306.
13. Berner ES, Webster GD, Shugerman AA, Jackson JR, Algina J, Baker AL, et al. Performance of four computer-based diagnostic systems. *N Engl J Med* 1994;330:1792-6.
14. Caratozzolo V. [Use of expert systems in internal medicine. Analysis of performance and reliability of the Quick Medical Reference]. [Italian] *Recenti Prog Med* 1995;86:492-5.
15. Ikechukwu A, Waseem A, Fox M, Barr C, Fisher K. Evaluation of Quick Medical Reference (QMR) as a teaching tool. *MD Comput* 1998;15:323-6.

Reprint requests to: Dr. Jane B. Lemaire, University of Calgary Health Sciences Centre, 3330 Hospital Dr. NW, Calgary AB T2N 4N1; fax 403 283-6151; lemaire@ucalgary.ca