6. Nussenblatt RB, Palestine AG, Chan CC. Cyclosporin-A therapy in the treatment of intraocular inflammatory disease resistant to systemic corticosteroids and cytotoxic agents. *Am J Ophthalmol* 1983; **93:** 275–81.
7. Palestine AG, Austin HA, Barlow JE, et al. Renal histopathologic alterations in patients treated with cyclosporin for uveitis. *N Engl J Med* 1986; **314:** 1293–98.
8. Towler HM, Whiting PH, Forrester JV. Combination low dose cyclosporin A and steroid therapy in chronic intraocular inflammation. *Eye* 1990; **4:** 514–20.
9. Shaw L. Critical issues in cyclosporin monitoring: report of the task force on cyclosporin monitoring. *Clin Chem* 1987; **33:** 1629–88.
10. Binder AI, Graham EM, Sanders MD, Dinning W, James WG, Denman AM. Cyclosporin A in the treatment of severe Behçet's disease. *Br J Rheumatol* 1987; **26:** 285–91.
11. Tabbara KF. Chlorambucil in Behçet's disease; a reappraisal. *Ophthalmology* 1983; **90:** 906–08.
12. Nussenblatt RB, Palestine AF, Chan CC, Mochizuki M, Yancey K. Effectiveness of cyclosporin therapy for Behçets disease. *Arthritis Rheum* 1985; **6:** 671–79.
13. Yazici H, Pazarli H, Barnes C, et al. A controlled trial of azathioprine in Behçet's syndrome. *N Engl J Med* 1990; **322:** 281–85.
14. De Kozak Y, Solkai J, Thilliaye B, Faure JP. S-antigen induced experimental autoimmune uveoretinitis in rats. *Curr Eye Res* 1981; **1:** 327–40.
15. Caspi RR, Roberge FG, McCallister CF, et al. T cell lines mediating experimental autoimmune uveoretinitis (EAU) in the rat. *J Immunol* 1986; **136:** 928–33.
16. Atalla L, Linker-Israeli M, Steinman L, Rao N. Inhibition of autoimmune uveitis by anti-CD4 antibody. *Invest Ophthalmol Vis Sci* 1990; **31:** 1264–70.
17. Wetzig R, Hooks JJ, Percopo CM, Nussenblatt R, Chan CC, Detrick B. Anti-Ia antibody diminishes ocular inflammation in experimental autoimmune uveitis. *Curr Eye Res* 1988; **7:** 809–18.
18. de Kozak Y, Mirshahi M, Boucheix C, Faure JP. Prevention of experimental autoimmune uveoretinitis by active immunization with autoantigen-specific monoclonal antibodies. *Eur J Immunol* 1987; **17:** 541–47.
19. Caspi RR, Kuwabana T, Nussenblatt RB. Characterization of a suppresor cell line which downgrades experimental autoimmune uveoretinitis in the rat. *J Immunol* 1988; **140:** 2579–84.
20. Beraud E, Kotake S, Chan CC, et al. T cell vaccination against experimental autoimmune uveitis (EAU). *Invest Ophthalmol Vis Sci* 1990; **31:** 213.
21. Roberge FG, Lorberboum-Galaski H, Le Hoang P, et al. Selective immunosuppression of activated T cells with the chimeric toxin IL-2-PE40. Inhibition of experimental autoimmune uveoretinitis. *J Immunol* 1989; **143:** 3498–502.
22. Nussenblatt RB, Caspi RR, Mahdi R, et al. Inhibition of S-antigen induced experimental autoimmune uveoretinitis by oral induction of tolerance with S-antigen. *J Immunol* 1990; **144:** 1689–95.
23. Qin S, Wise M, Cobbold S, et al. Induction of tolerance in peripheral T cells with monoclonal antibodies. *Eur J Immunol* 1990; **20:** 2737–45.
24. Kawashima H, Fujino Y, Mochizuki M. Effects of a new immunosuppressive agent FK506, in experimental uveoretinitis in rats. *Invest Ophthalmol Vis Sci* 1988; **29:** 1265–71.
25. Dumont FJ, Melino MR, Staruch MJ, et al. The immunosuppressive macrolides FK506 and rapamycin act as reciprocal antagonists in murine T cells. *J Immunol* 1990; **144:** 1418–24.

CLINICAL PRACTICE

Comparison of computer-aided and human review of general practitioners' management of hypertension

JOHAN VAN DER LEI MARK A. MUSEN EMIEL VAN DER DOES
ARIE J. MAN IN 'T VELD JAN H. VAN BEMMEL

Computer programs that automatically review decisions can help physicians provide better patient care. In the Netherlands, the ELIAS computer information system has replaced paper medical records in some general practices. We have written a computer program called 'HyperCritic' that audits general practitioners' management of patients with essential hypertension by taking patient-specific data from the ELIAS system. We investigated whether the computer-based medical records contain sufficient information to generate critiques, and compared the limitations of audit by hypercritic with those of review by a panel of eight physicians.

Hypercritic and the physicians independently reviewed the medical records of 20 randomly selected patients with hypertension and commented on the decisions made at each of 243 patient visits. Of 468 comments on patient management, 260 were judged correct by six or more of the physicians; hypercritic also made 118 of these 260 comments. The main reasons why the program did not produce the other 142 comments were: insufficient data in the computer-based medical record; absence of sufficient medical consensus; and omissions in the database of hypercritic. Calculation of an "index of merit" ([sensitivity + specificity] − 1) for individual reviewers showed that hypercritic performed better (index of merit 0·62) in its limited domain than did physician reviewers (0·3–0·56).

At least in hypertension management, automated review of computer-based medical records compares favourably with review by physicians. Further development of computer-aided clinical audit requires the introduction of computer-based medical records that capture the reasoning of physicians, and of widely accepted practice guidelines.

Lancet 1991; **338**: 1504–08.

Introduction

The storing of medical records on computer allows systematic analysis of past clinical experience, thus

ADDRESSES **Departments of Medical Informatics** (J van der Lei, MD, Prof J H. van Bemmel, PhD), **Institute for Primary Care** (Prof E van der Does, MD), **and Department of Internal Medicine I** (Prof A. J. Man in 't Veld, MD), **Faculty of Medicine and Health Sciences, Erasmus University, Rotterdam, Netherlands; and Section on Medical Informatics, Stanford University School of Medicine, Stanford, California, USA** (M A Musen, MD) Correspondence to Dr J. van der Lei, Department of Medical Informatics, Faculty of Medicine and Health Sciences, Erasmus University, PO Box 1738, 3000 DR Rotterdam, Netherlands

providing guidelines for future practice and policies.[1] Programs to support clinical decision making[2] have been developed in parallel with computer-stored medical records. However, despite early optimism about the future of computer-assisted decision making,[3,4] these programs have not come into widespread use. Miller and Masarie[5] argue that one of the fundamental causes of this lack of success is the so-called "Greek oracle" model for decision support incorporated in early systems. In this model, the physician, unable to resolve a given medical problem, submits all relevant patient information to the system. The physician is then a passive observer who waits for the oracle to solve the problem.

An alternative to the Greek oracle model is the "critiquing" model. Here, the physician, in addition to submitting patient-specific data, lets the system know the decisions he intends to make. The system evaluates these decisions and expresses agreement or suggests alternatives.[6] Programs based on the critiquing model for computer-aided decision support can be integrated with computer-based medical records. Physicians enter data into a computer and, behind the scenes, the medical-records system forwards patient data to a critiquing program. The critiquing program evaluates the decisions of physicians and, if inappropriate, suggests reasoned alternatives.

In the Netherlands, the ELIAS program is used by general practitioners (GPs) to replace paper-based medical records with computer-based records.[7] We used data taken from these computer-based medical records on GPs' decisions about management of patients with hypertension to compare the performance of human reviewers with that of a critiquing program called 'HyperCritic'.[8]

Patients and methods

Medical records program

In the Netherlands, the ELIAS program is used at over 300 sites by more than 500 GPs (10% of all Dutch GPs). The GP enters patient data into ELIAS via a video-display unit and keyboard in the consulting room. At some sites, ELIAS is electronically connected to computers at hospitals and laboratories in the region. Thus, laboratory reports, for example, can be sent directly from remote computers. Laboratory data, measurements obtained during physical examination, and prescription information are always entered in a machine-interpretable format, but not all data in ELIAS are coded in this way. One of two diagnosis-coding schemes can be used by ELIAS.[9,10]

Critiquing program

We have developed the hypercritic program, which audits a physician's treatment of hypertensive patients.[8] Hypercritic takes as input the electronic ELIAS medical record for a single patient; it will not request any data other than those available in this medical record. The output of hypercritic is in the form of text. Hypercritic does not reside on the same computer as ELIAS, but the two programs are electronically linked. The program has not yet been released for routine use.

Hypercritic offers comments on drug therapy and laboratory tests given to patients with benign essential hypertension. Hypercritic does not attempt to use data supplied by ELIAS to assign the diagnosis of hypertension; rather, the program audits a physician's therapy of hypertension only after the physician has made that diagnosis. Furthermore, hypercritic does not judge any diagnostic investigation as inappropriate based solely on the absence of an apparent indication. The system will, however, report when a test that is apparently indicated is not done.

Hypercritic generates comments in two stages. First, the program interprets the medical record to discover the actions of the GP at a given patient visit (eg, starting a new drug, continuing treatment with a drug, or replacing one drug with another). Second, to review each action, the program: (1) searches the medical record for conditions that contraindicate the action (eg, contraindications to specific drugs); (2) determines whether preparatory procedures required before the action have been done (eg, the evaluation of kidney function before starting treatment with an angiotensin-converting-enzyme inhibitor); (3) determines whether the GP has done the routine monitoring required by the action (eg, monitoring potassium when continuing treatment with thiazides); and (4) searches for any undesirable consequences of the action (eg, drug side-effects). Reviewing computer-stored medical record requires specific medical knowledge; therefore, hypercritic contains drug information (eg, normal doses, contraindications, side-effects, interactions), diagnostic workup requirements for hypertensive patients, and criteria for judging the efficacy of the treatment. The program has been described in detail elsewhere.[8]

TABLE I—PART OF THE ELIAS MEDICAL RECORD OF ONE PATIENT

Visit date	Type of record	Record
May 29, 1987	Subjective	*Visit for monitoring blood pressure*
	Objective	BP 135/100 mm Hg, pulse 76 bpm, weight 69·5 kg
	Plan	Capoten 25 mg TDS
Aug 20, 1987	Subjective	*Visit for monitoring blood pressure*
	Objective	BP 160/100, pulse 72
	Plan	Capoten 50 mg TDS, moduretic OD
Jan 20, 1988	Subjective	*Visit for monitoring blood pressure*
	Objective	BP 135/85, pulse 72
	Plan	Capoten 50 mg TDS, moduretic OD
April 8, 1988	Subjective	*Visit for monitoring blood pressure*
	Objective	BP 140/80, pulse 72
	Plan	Capoten 50 mg TDS, moduretic OD
May 26, 1988	Subjective	*Rash on forearms*
	Assessment	Photosensitivity (3660); *possibly caused by capoten*
	Plan	Continue treatment. *Request lab*
June 3, 1988	Subjective	*Still rash on forearms. Not all parts that have been exposed to sunlight have been affected*
	Objective	Creat 138 μmol/l, BUN 9·4 mmol/l, ESR 6 mm/h, HCT 0·45, Hb 8·8 mmol/l, LDH 219 U/l, SGOT 14 U/l, SGPT 9 U/l, Alk phos 39 U/l, Leuco count 7·5 /nl; diff: eos 3, segs 68, lymphs 24, monos 5. Ery count 4·92 /pl. Ery: no abnormalities
	Assessment	*Photosensitivity less probable. Not all parts affected. Has already used capoten for a long period.*
	Plan	Continue treatment

Data that were not coded (ie, free text) are shown in italics 'Capoten' (Squibb) is captopril 'Moduretic' (Merck Sharp & Dohme) is a preparation containing hydrochlorothiazide and amiloride

Patients

For this study, we selected the two oldest ELIAS installations (installed in 1985 and 1986) as a source of patient data. The systems are used in two group practices by eight GPs who provide primary care for about 13 000 patients. The practices no longer maintain paper medical records, since GPs enter clinical data directly into the computer at the time of each consultation.

Patients met the following eligibility criteria: (1) hypertension had been diagnosed by the GP; (2) ELIAS was in use when hypertension was first diagnosed; (3) the patient had not been seen by a hospital-based physician for evaluation of hypertension; and (4) the patient was not pregnant. During the year after the introduction of ELIAS, 83 patients who met the eligibility criteria were seen at the study clinics. We randomly selected 20 of the patients, and collected data on the patients recorded by ELIAS from its introduction until 1989. The medical records of the 20 patients during this period described 243 visits covering 44 patient-years of hypertension management. The patients made between 4 and 26 visits (median 11) to their GPs during the study period. Patients were aged 21 to 80 (median 61) years; 8 patients were female and 12 male.

TABLE II—COMMENTS GENERATED BY REVIEW OF MEDICAL RECORD SHOWN IN TABLE I

Visit date	Comment	No of physicians who agreed with the comment	Did hypercritic generate comment?
Aug 20, 1987	The initial dose of hydrochlorothiazide is too high (1 tablet moduretic contains 50 mg hydrochlorothiazide and 5 mg amiloride).	6	Yes
	The increment in the dose of captopril (from 75 to 150 mg) is too great.	2	Yes
	The dose of captopril is high already (75 mg); increasing the dose to 150 is useless.	5	No
	Do not combine a potassium-sparing diuretic with an ACE-inhibitor.	5	Yes
	Due to the treatment with the diuretic, measure the potassium.	7	Yes
	The blood-pressure level is acceptable; do not increase the drug regimen.	2	No
Jan 20, 1988	The interval between this visit and the previous visit is too long. On the previous visit, the drug regimen was changed; the effect should be measured within 6 weeks.	8	Yes
April 8, 1988	Blood pressure has responded well to treatment; reduce drug dose.	8	No
May 26, 1988	Photosensitivity could be caused by captopril; discontinue treatment with captopril.	4	Yes
	Photosensitivity could be an allergic reaction; discontinue all drugs.	2	No
	When treating a patient with antihypertensive drugs, measure the blood pressure at each visit.	5	Yes
June 3, 1988	The creatinine has increased (138 μmol/l). Captopril may cause a decrease in kidney function; discontinue captopril.	5	Yes
	Potassium-sparing diuretics are contraindicated because of the increased creatinine (138 μmol/l).	5	Yes

Review of patients' records

Printouts of each patient's medical records were submitted to eight physicians for review. Two of the eight (physicians I and II) were experienced cardiovascular physicians who worked in university hospitals in different parts of the Netherlands. Three (physicians III–V) were GPs working in university institutes for primary care who had an interest in the treatment of cardiovascular diseases in primary care. These five physicians nominated practising GPs who had no medical school position, from whom we randomly selected three GPs (physicians VI–VIII).

We randomised the order in which each physician reviewed the patients' records. Table I shows part of one of the medical records that was reviewed. Working independently, reviewers indicated on a questionnaire whether each examination or therapeutic action was appropriate or inappropriate, and whether any tests or interventions were missing from a patient's record. When the reviewers deemed an item to be inappropriate or missing, they justified their assessment. Reviewers also made written comments on each patient visit.

For each visit, we compiled a list of reviewers' comments. We considered a comment to be an individual remark about a specific action (or absence of an action) described in the medical record. For example, "I would not treat this patient with this drug; if you insist on treating the patient with this drug, then the dosage is too high" is two comments.

Patients' records were reviewed also by the hypercritic program. We matched the comments of the computer system to the comments of the eight reviewing physicians. Three independent referees verified this matching. Not all hypercritic's comments matched those of reviewing physicians; these extra comments were added to the physicians' comments.

In a subsequent "Delphi-type" round,[11] we gave the reviewing physicians the list of comments for each visit that had been generated by the reviewers and hypercritic, and asked them to state whether they believed each comment to be correct. Table II shows the comments based on the ELIAS medical record shown in table I.

Analysis

We used two different peer-review standards for judging performance. First, we compared the output of hypercritic with the combined opinion of the reviewing physicians. A comment judged to be correct by six or more of the eight reviewing physicians was accepted; we assumed that such a comment was both valid and clinically relevant. For each accepted comment not made by hypercritic, we established the reason for the omission. Comments outside the predefined domain of the program were labelled "outside domain". A comment inside the program's domain which required the interpretation of free text (ie, uncoded data), was labelled "free-text dependent". When the comment was inside the program's domain and did not require interpretation of free text, we determined whether the program had considered the comment but had not made it because a threshold in the program had not been passed. Such comments were labelled "considered", and we identified the threshold that, if changed, would cause hypercritic to produce that comment. Finally, comments that did not fall into any of the above categories were labelled "omission", and we identified the knowledge required by hypercritic to generate the comments. Thus, whereas "outside domain" represents our deliberate decision to limit the scope of the program, "omission" represents our failure to include in the program information that should have been present.

The second assessment method compared individual sources of comments (individual reviewers and hypercritic) with each other. Outside-domain and free-text-dependent comments were removed from the total set of comments. For the remaining comments, we assumed that the failure of hypercritic to generate a given comment was equivalent to a reviewing physician judging that comment incorrect. With this assumption, kappa statistics[12] were used to assess agreement among all reviewers, including hypercritic. Although a kappa value provides a measure of the degree of agreement among reviewers, the nature of any disagreement is not clarified. We therefore assumed that the majority opinion (ie, five or more) of the reviewers (hypercritic included) correctly determined in each case whether a given comment was appropriate. With this additional assumption, various indices of efficacy could be estimated. The sensitivity of a given reviewer was defined as the fraction of comments judged correct by the majority of reviewers that the individual reviewer also judged to be correct, and the specificity was defined as the fraction of incorrect comments that the individual reviewer also judged to be incorrect. The predictive value of a positive judgment was the fraction of comments that the individual reviewer judged to be correct that the majority of reviewers also judged to be correct, and the predictive value of a negative judgment was the fraction of comments that the individual

TABLE III—COMMENTS ACCEPTED BY PHYSICIANS BUT NOT REPRODUCED BY HYPERCRITIC

—	Type of comment			
	Diagnosis of hypertension	Selection of optimum treatment	Execution of treatment	Total
No accepted	139	52	69	260
Not reproduced by hypercritic				
Total	93	31	18	142
Outside domain	80	0	0	80
Free-text dependent	5	23	1	29
Considered	6	7	9	22
Omission	2	1	8	11

TABLE IV—PERFORMANCE OF REVIEWERS

Reviewer	Sensitivity	Specificity	PV positive	PV negative	Index of merit
Physician					
I	0·94	0·36	0·79	0·70	0·30
II	0·86	0·70	0·88	0·66	0·56
III	0·72	0·82	0·91	0·54	0·54
IV	0·65	0·75	0·87	0·46	0·40
V	0·73	0·69	0·86	0·50	0·42
VI	0·70	0·78	0·89	0·50	0·48
VII	0·88	0·52	0·82	0·63	0·40
VIII	0·74	0·77	0·89	0·54	0·51
Hypercritic	0·74	0·88	0·94	0·57	0·62

PV = predictive value.

reviewer judged to be incorrect that the majority of reviewers also judged to be incorrect. Index of merit was defined as (sensitivity + specificity) − 1; thus, this index ranges from −1 to +1.

Results

468 comments were made, a mean of 1·9 per patient visit. The number of comments per patient record ranged from 7 to 70 (median 19). Of the 468 comments, 169 were also made by hypercritic.

260 of the 468 comments were judged correct by six to eight physicians. Of these 260 accepted comments, hypercritic failed to reproduce 142 (table III). 2 comments on diagnosis were not made because of omissions in the database: in both cases, the program failed to detect a possible primary cause of the hypertension. 7 of the accepted comments on selection of therapy that hypercritic failed to reproduce were classified as considered: 5 were about the role of angiotensin-converting-enzyme inhibitors as a drug of first choice, and 2 about whether blood pressure was sufficiently high to warrant drug treatment. Hypercritic's 1 omission in the selection-of-therapy category was when it failed to recommend stopping oral contraceptives in a patient with hypertension. Of the 69 accepted comments on execution of therapy, the program failed to reproduce 18. In 1 case, hypercritic failed to detect a drug side-effect (depression caused by a beta-blocker) because the patient's symptoms were recorded only in free text; if the GP had coded the data, hypercritic would have noted the possible side-effect. Another 9 comments were considered by hypercritic: 8 of these dealt with reducing the dose of antihypertensive drugs once the blood pressure had fallen; the peer reviewers advised an earlier reduction in drug dosage than the program would have recommended. The program failed to make 8 comments because of omissions in its database: in 7 cases, the cause was omission of criteria specifying how blood pressure should be monitored after drug treatment has been discontinued, and in the remaining case hypercritic failed to recognise a side-effect because it was not part of the program's database.

Before further analysis, comments classified as outside the domain of hypercritic or free-text-dependent were removed from the total set of comments, leaving 298. Kappa values for pairs of physician reviewers ranged from 0·09 to 0·45, and for the pairing of any physician reviewer with hypercritic from 0·08 to 0·46. Physician I had the highest sensitivity (true-positive rate), but the lowest specificity (true-negative rate). Hypercritic had the highest specificity, and the highest predictive value of a critique (table IV). The physicians' index of merit was between 0·3 and 0·56, whereas hypercritic had an index of merit of 0·62.

Discussion

The barriers to widespread adoption of computer-aided decision support in medicine are substantial.[13] One of the most important barriers is the inability of present computer programs to acquire clinically relevant data automatically from hospital and office information systems.[14] The hypercritic program was written in an attempt to overcome this difficulty. This study was designed to investigate whether the data available from the ELIAS computer records were sufficient to generate clinically useful comments, and to validate the hypercritic program as a reviewer of GPs' decision making. Although we did not examine whether physicians change their behaviour in response to the program's comments, other researchers have shown important effects of computer-generated advice on physicians' actions.[15,16]

The ELIAS medical records contained sufficient information for human reviewers and for hypercritic to generate substantial critiques. The number of comments generated suggests deficiencies in the management of hypertensive patients, and that human reviewers and hypercritic can provide comments that are useful to physicians.

Hypercritic has difficulty in determining the diagnostic and treatment goals of the physician. Knowledge of those goals is essential for assessment of the physician's actions. Medical records contain data describing the patient's state (eg, the results of laboratory tests) and the objectives of the treating physician (eg, a list of treatment goals). But the medical record mainly records what was done rather than why. Moreover, physicians do not record all actions taken or all decisions made. One of the goals of the problem-oriented medical record[17] is to facilitate mutual understanding among physicians by introducing the notion of a "problem" as one of the axes along which to structure the medical data. By assignment of the entries in the medical record to the problems that the physician has identified, some of the ambiguities in those entries can be clarified.

Data in the computer-based medical record that describe the reasoning of the physician are typically in the form of free text and cannot be linked to other data in the record. The absence of coded data that might allow the computer to ascertain the intentions of the treating physician is a fundamental obstacle to computer-based audit. Hypercritic does not judge a diagnostic investigation to be inappropriate based solely on the absence of an apparent indication. Since hypercritic has knowledge limited to hypertension management and access only to coded data in the computer medical record, it is unable to make the overall assessment of the patient necessary for such a judgment—the program reviews only bits and pieces of the care that a patient receives.

The success of computer-based audit is also limited by the availability of medical knowledge. The low kappa values between different physician reviewers emphasises the fact that many decisions made by physicians are arbitrarily variable, and that this arbitrariness represents, for at least some patients, suboptimum care.[18] In much of medicine, there is no consensus that defines proper therapy. Yet audit requires a set of criteria. The development of an auditing program that uses the critiquing model forces the explicit definition of those criteria. McDonald et al[19] reported that computer-based review had a significant effect on care only when physicians agreed beforehand the ideal approach to a problem.

In the absence of an absolute standard of care, hypercritic embodies just one set of opinions from a range of possible choices. The system thus contains one set of criteria for reaching particular conclusions, whereas some reviewing physicians apparently used other sets of criteria. In our study, hypercritic tended to be more lenient (that is, less critical) than the reviewing physicians. The low sensitivity of hypercritic (table IV) shows its leniency.

Hypercritic provides decision support as a byproduct of routine data-management; the physician does not specifically ask for advice. Therefore, in writing the program, we had to avoid generating excessive numbers of unwanted comments. The critiques produced by hypercritic, and similar programs, must have high predictive value. However, increasing the sensitivity of the computer system will probably reduce its predictive value by causing the generation of more false-positive comments.

Pressure from legislative bodies, insurers, peer-review organisations, hospitals, and physicians and patients may lead to increasing use of automated review of medical records. As such technology is put into place, we must remember that computer systems are based on models, and that the models are limited in what they contain. There is no evidence that the capabilities of computers will ever approach those of human beings in dealing with unexpected circumstances, in understanding patients in their social context, in integrating the often complex and confounding presentation of a disease into a coherent pattern, or in dealing with ethical issues.[20] Program writers may choose not to model certain aspects of medical care, and to inform users of these limitations in the programs. Before any computer-based decision-support program is released, a formal evaluation with real patient data must be done so that the limitations of the system will be discovered.

We thank Prof R. van Strik and Dr H. J. A. Schouten for their participation in designing this study and subsequent data analysis. We also thank those GPs who made their medical records available for scrutiny, and the physicians who participated in this study. The study was supported by the Netherlands Heart Foundation (grant 88.236). M. A. M. is the recipient of FIRST Award LM05157 from the United States National Library of Medicine.

REFERENCES

1. McDonald CJ, Tierney WM. Computer-stored medical records: their future role in medical practice. *JAMA* 1988; **259:** 3433–40.
2. Shortliffe EH. Computer programs to support clinical decision making. *JAMA* 1987; **258:** 61–66.
3. Yu VL, Fagan LM, Wraith SM, et al. Antimicrobial selection by a computer: a blinded evaluation by infectious disease experts. *JAMA* 1979; **242:** 1279–82.
4. Miller RA, Pople HE Jr, Myers JD. Internist-1: an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982; **307:** 468–76.
5. Miller RA, Masarie FE. The demise of the Greek oracle model for medical diagnosis systems. *Methods Inf Med* 1990; **29:** 1–2.
6. Miller PL. Expert critiquing systems: practice-based medical consultation by computer. New York: Springer-Verlag, 1986.
7. Westerhof HP, Boon WM, Cromme PVM, et al. ELIAS: support of the Dutch general practitioner. In: Reichertz PL, Engelbrecht R, Picollo U, eds. Present status of computer support in ambulatory care. New York: Springer-Verlag, 1987: 1-10.
8. Van der Lei J, Musen MA. A model for critiquing based on automated medical records. *Comput Biomed Res* 1991; **24:** 344–78.
9. Lamberts H, Wood M, eds. International classification of primary care. Oxford: Oxford University Press, 1987.
10. Royal College of General Practitioners. Classification of diseases, problems and procedures. London: Royal College of General Practitioners, Occasional Paper 26, 1984.
11. Dalkey NC. An experimental study of group opinion: the Delphi method. *Futures* 1969: 408–26.
12. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; **20:** 37–46.
13. Shortliffe EH. Testing reality: the introduction of decision-support technology for physicians. *Methods Inf Med* 1989; **28:** 1-5.
14. Cooper GF, Musen MA. The 1990 AAAI spring symposium on artificial intelligence in medicine. *Artif Intel Med* 1990; **2:** 293–98.
15. Evans RS, Larsen RA, Burke JP, et al. Computer surveillance of hospital-acquired infections and antibiotic use. *JAMA* 1986; **256:** 1007–11.
16. McDonald CJ. Protocol-based computer reminders: the quality of care and the non-perfectibility of man. *N Engl J Med* 1976; **295:** 1351–54.
17. Weed LL. Medical records, medical education, and patient: the problem-oriented record as a basic tool. Cleveland, USA: Case Western Reserve Press, 1971.
18. Eddy DM. Clinical decision making: from theory to practice. The challenge. *JAMA* 1990; **263:** 287–90.
19. McDonald CJ, Hui SL, Smith DM, et al. Reminders to physicians from an introspective computer medical record. *Ann Intern Med* 1984; **100:** 130–80.
20. Blois MS. Clinical judgment and computers. *N Engl J Med* 1980; **303:** 192–97.

HYPOTHESIS

Achlorhydria and gastric carcinogenesis

JOHN P. SEERY

The association between achlorhydria and gastric carcinogenesis can be explained by a simple hypothesis based on the known cytological properties of neoplastic cells and the physiology of the stomach. Normal gastric secretions might ensure the rapid elimination of carcinomatous cells. Achlorhydria could be an important permissive factor in the development of gastric carcinoma.

Lancet 1991; **338:** 1508–09.

The association between achlorhydria and gastric carcinoma is well known. Prospective studies have shown a four to six fold rise in the relative risk for gastric carcinoma in achlorhydric patients.[1] Furthermore, over 60% of patients with gastric cancer have achlorhydria compared with an incidence of only 20% in normal individuals of a similar age.[2]

Achlorhydria may predispose to the proliferation of bacteria in the stomach; these bacteria convert nitrites to nitrosamines and nitrosamides.[3] Many such N-nitroso compounds are powerful carcinogens and are thought to act on the gastric epithelium. Recent studies have shown that certain bacterial isolates can lead to formation of N-nitroso compounds, but the low rates of reaction make their clinical relevance doubtful.[4] Moreover, the rate of acid-catalysed N-nitroso compound formation decreases with increasing pH.[5] Whether increased bacterial synthesis of nitrosamines and nitrosamides can compensate for this fall remains open to question.[6] The role of increased N-nitroso compound concentrations in linking achlorhydria with gastric carcinogenesis is, therefore, unproven. Achlorhydria might render the gastric mucosa more susceptible to the action of other carcinogenic agents.

ADDRESS. **Department of Clinical Pharmacology, Royal Infirmary of Edinburgh, Lauriston Place, Edinburgh EH3 9YW, UK.** Correspondence to Dr J P Seery, MB.