Automated attention flags in chronic disease care planning

James R Warren, Joseph T Noone, Brian J Smith, Richard Ruffin, Peter Frith,
Berend J van der Zwaag, Gleb V Beliakov, Heath K Frankel and Heather J McElroy

LEAPE DEFINES *practice guidelines* as standardised specifications for care developed by a formal process that incorporates the best scientific evidence of effectiveness with opinions of experts in the fields.[1] Wide variations in clinical practice exist,[2] not all of which are "best practice". Guidelines have been developed to improve patient care and reduce cost and variations in clinical practice.[3] When successfully implemented, practice guidelines improve health outcomes.[4]

Guideline implementation is most likely to be effective when patient-specific advice is provided during a consultation,[4-5] and, in particular, "when the guideline is made accessible through computer-based, patient-specific reminders that are integrated into the clinician's workflow".[6]

Our study addressed the utility of further decision support for general practitioners (beyond project-specific, paper guidelines) in chronic disease care planning. We considered a computer-based decision tool, Care Plan On-Line (CPOL), formulated to provide attention flags in accordance with project-specific guidelines of the South Australian HealthPlus Coordinated Care Trial. The potential value of automated decision support varies with the level of agreement between the guidelines and the (unsupported) decisions of the care coordinators (less useful if high agree-

ABSTRACT

Objectives: To assess the value of computerised decision support in the management of chronic respiratory disease by comparing agreement between three respiratory specialists, general practitioners (care coordinators), and decision support software.

Methods: Care guidelines for two chronic obstructive pulmonary disease projects of the SA HealthPlus Coordinated Care Trial were formulated. Decision support software, Care Plan On-Line (CPOL), was created to represent the intent of these guidelines via automated attention flags to appear in patients' electronic medical records. For a random sample of 20 patients with care plans, decisions about the use of nine additional services (eg, smoking cessation, pneumococcal vaccination) were compared between the respiratory specialists, the patients' GPs and the CPOL attention flags.

Results: Agreement among the specialists was at the lower end of moderate (intraclass correlation coefficient [ICC], 0.48; 95% CI, 0.39–0.56), with a 20% rate of contradictory decisions. Agreement with recommendations of specialists was moderate to poor for GPs (κ, 0.49; 95% CI, 0.33–0.66) and moderate to good for CPOL (κ, 0.72; 95% CI, 0.55–0.90). CPOL agreement with GPs was moderate to poor (κ, 0.41; 95% CI, 0.24–0.58). GPs were less likely than specialists or CPOL to decide in favour of an additional service ($P < 0.001$). CPOL was 87% accurate as an indicator of specialist decisions. It gave a 16% false-positive rate according to specialist decisions, and flagged 61% of decisions where GPs said No and specialists said Yes.

Conclusions: Automated decision support may provide GPs with improved access to the intent of guidelines; however, further investigation is required.

MJA 2001; 175: 308-312

ment). Automated decision support must also match expert interpretation of the guidelines in specific cases. Thus, we analysed agreement of specialists, care coordinators and CPOL on respiratory care planning decisions.

Advanced Computing Research Centre, University of South Australia, Adelaide, SA.
James R Warren, BSc, PhD, Associate Professor; Joseph T Noone, BHSc(Hons), PhD Student.
North Western Adelaide Health Service, Adelaide, SA.
Brian J Smith, MBBS, PhD, Respiratory Physician; Richard Ruffin, MBBS, PhD, Head, Division of Medicine; Heather J McElroy, BSc(Hons), Biostatistician.
Repatriation Hospital, Adelaide, SA.
Peter Frith, MBBS, PhD, Head, Department of Respiratory Medicine.
EUREGIO Computational Intelligence Centre, University of Twente, Enschede, The Netherlands.
Berend J van der Zwaag, MSc, Research Assistant.
School of Computing and Mathematics, Deakin University, Melbourne, VIC.
Gleb V Beliakov, MSc, PhD, Lecturer.
Medical Communications Associates Pty Ltd, Belair, SA.
Heath K Frankel, BSc(Hons), Intranet Developer.
Reprints will not be available from the authors. Correspondence: Associate Professor J R Warren, Advanced Computing Research Centre, University of South Australia, Adelaide, SA.
warren@cs.unisa.edu.au

METHODS

Box 1 illustrates the timeline and document flow relevant to our study.

SA HealthPlus trial

The SA HealthPlus trial ran from July 1997 to December 1999 as a First Round Coordinated Care Trial.[7]

SA HealthPlus included 10 disease-specific projects. Project-specific materials were compiled by Care Mentor groups, which comprised specialists, GP opinion leaders, nurses, social workers and consumer advocates. The three authors who are respiratory physicians (BJS, RR, PF) had a prominent role in developing evidence-based guidelines that formed the basis of the SA HealthPlus respiratory care planning guidelines.

The target group for SA HealthPlus was people with complex health and social problems who required multiple

services, including people with diabetes, respiratory, cardiac or mental illnesses. SA HealthPlus admission criteria included multiple hospital emergency events in the past year, a diagnosis appropriate to the project, and willingness to participate.

Each patient nominated a GP to act as his or her care coordinator. With the support of a service coordinator (usually a nurse), a project-specific initial medical assessment was made and a Problems and Goals statement in the patient's own words was collected. The care coordinator then designed a 12-month care plan for the patient (Box 1). A single care planning form gave the recommended "A level" services for each of mild, moderate and severe patient levels (determined in an objective way from the initial medical assessment). The care coordinators then used their judgment, supplemented by a project-specific guideline booklet, to select additional "B level" services to adapt the base care plan to individual patient needs.

Software

An intranet-based system (running on the South Australian Government's Statenet) was devised. This system, Care Plan On-Line (CPOL), is designed to allow service coordinators and care coordinators to remotely view and update electronic patient records (including care plans, initial and ongoing medical assessments, and Problems and Goals statements) of SA HealthPlus patients. This software also provides access to the project-specific guidelines.

Electronic patient records and guidelines are integrated in CPOL via dynamic decision-support flags (a yellow "?" or a red "!") that appear next to relevant signs on the patient record display (eg, the dietitian service may be flagged next to an overly low body mass index) and adjacent to services on the list from which the care coordinator specifies the patient's care plan. The criteria for flag display are based on formalisation of the guidelines in so far as they could be readily automated from the SA HealthPlus electronic patient record. The purpose of the flags is to attract the attention of the care coordinator, who can display the related care guideline and review the evidence for taking an action. The care coordinator can display a guideline in the

1: Study timeline and document flow in the SA HealthPlus Coordinated Care Trial

Shaded boxes indicate the documents compared in this study. CPOL = Care Plan On-Line
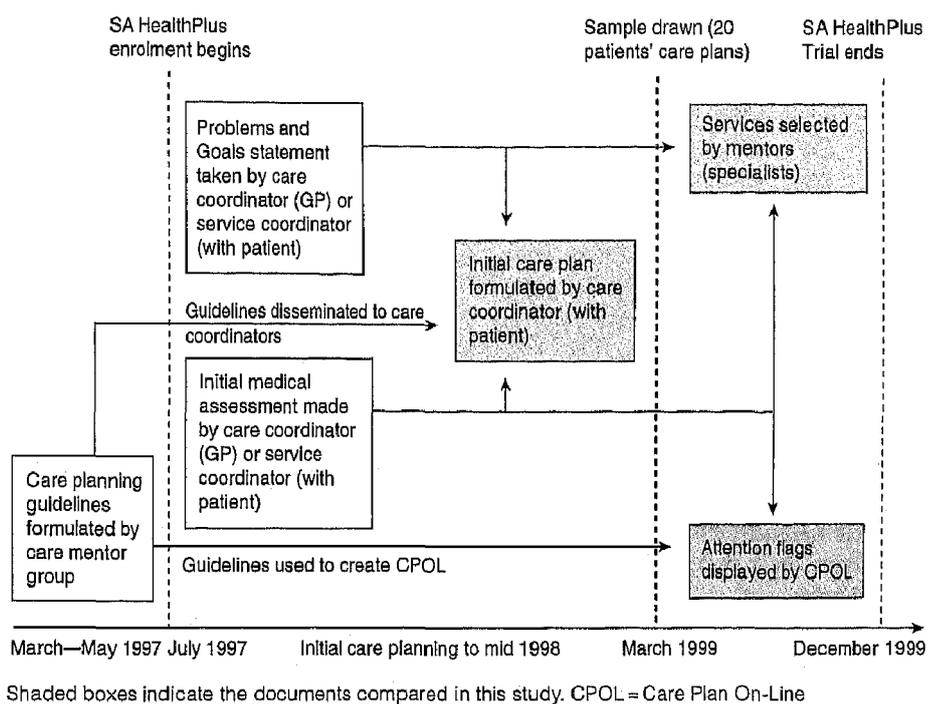
2: Summary of the additional services included in the study

Service	Service type*	Quality of evidence†	Care coordinator options	Care Plan On-Line options§
Counselling on diet/ weight/alcohol/ exercise	Informal GP activity	I	Not recorded	Attention flag displayed: Yes or No
Dietitian	B level	II	Included in Care Plan: Yes or No	Attention flag displayed: Yes or No
Influenza vaccination	Provisional-A level	I	Included in Care Plan: Yes or No	Attention flag displayed: Yes or No
Pneumococcal vaccination	Provisional-A level	II	Included in Care Plan: Yes or No	Attention flag displayed: Yes or No
Bone density screening	B level	III	Included in Care Plan: Yes or No	Attention flag displayed: Yes or No
ECG to look for pulmonary hypertension	B level	II	Included in Care Plan: Yes or No	Attention flag displayed: Yes, No or "no triggers"¶
Smoking cessation	Provisional-A level	I	Included in Care Plan: Yes or No	Attention flag displayed: Yes or No
Physiotherapy	B level	I	Included in Care Plan: Yes or No	None: "no triggers"**
Pharmacy medication management	B level	III	Not universally available‡	Attention flag displayed: Yes or No

* B-level services could be included in the care plan if the care coordinator thought them appropriate. Provisional-A-level services were considered mandatory if indications were present and there were no contraindications. †Categories of evidence:[9] (I) Based on well designed randomised controlled trials, meta-analyses, or systematic reviews; (II) Based on well designed cohort or case–control studies; (III) Based on uncontrolled studies or consensus. ‡Pharmacy Medication Management (formal review of a patient's medications) was piloted as a component of SA HealthPlus after the trial began. § In this analysis we make no distinction between the "?" and "!" flags in CPOL. ¶ For 9 of the 20 patients, the observations CPOL uses for automatic flagging of ECG were left blank on the initial medical assessment form. ** The version of CPOL used for this analysis had no automated triggers to flag physiotherapy; moreover, the initial medical assessment form was very scant on explicit observations that would indicate physiotherapy.

context of the present patient by using the flag as a hypertext link, or the guidelines can be reviewed ad hoc through an index.[8]

Subjects

Two of the SA HealthPlus projects focused on chronic obstructive pulmonary disease, and involved about 400 patients from metropolitan Adelaide. For our study, 20 patients were randomly selected from the enrolments of these two projects (10 from each).

Protocol

Specialist (mentor) decisions: For each patient, initial medical assessment forms and Problems and Goals statements were copied from their paper files, de-identified, and provided to the three respiratory specialist authors ("mentors"). The initial 12-month care plans formulated manually (ie, without the use of CPOL) by the patients' care coordinators were also retrieved, but not distributed to the mentors. The three mentors were asked to assess and recommend independently each of nine optional additional services (Box 2) for inclusion in the care plans. Each mentor provided one of three opinions — Yes, No, or Insufficient Information — for each service as applicable to each patient.

With each mentor having three response options, there were 10 possible combinations of responses for each service. We reduced these combinations to an aggregate mentor decision of Yes, No, Insufficient Information, or Contradic-

tory (at least one specialist says Yes and one says No).

CPOL flags: The patients' initial medical assessment information was entered into CPOL to cause display of automated attention flags in accordance with CPOL's model of the project guidelines (Box 2). This was viewed by the researchers; GPs did not have access to CPOL during the trial.

Evaluation

To evaluate the level of agreement of mentors, we used intraclass correlation for each service and overall. Pairwise agreement between mentors was also examined using intraclass correlation.[10] Decisions were coded as 0 for No, 0.5 for Insufficient Information and 1 for Yes.

To evaluate the level of agreement between mentors and care coordinators, and between mentors and CPOL, we amalgamated the three mentors' decisions to a single Yes/No, omitting cases where the mentors disagreed or where all indicated Insufficient Information. Agreement was examined with the κ statistic. We compared the proportion of affirmative decisions for mentors and care coordinators using McNemar's paired proportion test.

In testing agreement, we take intraclass correlation coefficients (ICCs[10]) or κ levels of 0.75 or higher as "relatively good" agreement and ICC or κ below 0.4 as "poor" agreement.[10,11] We term ICC or κ between 0.4 and 0.75 as "moderate" agreement.

We also provide a descriptive analysis of how CPOL flags distribute over mentor decisions and of the joint distribution of CPOL flags, mentor decisions and care coordinator decisions.

RESULTS

Agreement between the three mentors over all services was at the lower end of moderate (ICC, 0.48; 95% CI, 0.39–0.56), but varied by service from no agreement beyond chance to perfect agreement (Box 3). The ICCs between mentor pairs over all services were 0.44, 0.44, and 0.54. Moreover, there was considerable mentor disagreement (20% rate of contradiction).

There was moderate to poor agreement between mentors and care coordinators (κ, 0.49; 95% CI, 0.33–0.66), with mentors more likely to recommend action (Box 4). Agreement between mentors and CPOL was moderate to good (κ, 0.72; 95% CI, 0.55–0.90) (Box 4).

In 112 decisions, mentors recommended services in 26 (23%; 95% CI, 15%–32%; $P<0.001$) more cases than care coordinators. Care coordinators almost never exceeded the aggregate mentor decision (2 cases in 112 decisions), but omitted 28 of the 69 services the mentors would have included in the care plan (Box 4).

Box 5 shows three-way analysis of mentor, care coordinator and CPOL recommendations. CPOL provided an indicator of mentor-positive recommendation of a decision in 17 of the 20 decisions within its design specification

3: Levels of agreement between the three mentors about the use of each additional service

| | Yes | | | Contradictory | | | Insufficient | No | | | | |
	YYY	YYI	YII	YYN	YIN	YNN	Information	IIN	INN	NNN	ICC	95% CI
Counselling on diet/weight/alcohol/ exercise	4	4		3	4		1	4			0.10	0.00–0.38
Dietitian	5	1	1	1	2		1	4	5		0.47†	0.21–0.74
Influenza vaccination	18			1						1	0.74*	0.58–0.91
Pneumococcal vaccination	1	8	8	2						1	0.12	0.00–0.40
Bone density screening	1	3	3	3				5	5		0.26	0.00–0.55
ECG to look for pulmonary hypertension		2	4	7	1	4			1	1	0.00	0.00–0.26
Smoking cessation	4									16	1.00*	
Physiotherapy		8	2	1	2		3	3	1		0.12	0.00–0.40
Pharmacy medication management	9			4	1				1	5	0.64*	0.43–0.85
Overall	42	26	18	17	14	5	5	16	13	24	0.48†	0.39–0.56

Mentors' decisions are summarised by Y for yes, N for no and I for insufficient information. For example, YYI indicates two specialists decided in favour of the action and one felt there was insufficient information. * Relatively good (or perfect) agreement. † Moderate agreement. ICC = intraclass correlation coefficient.

where care coordinators decided against the service (and 17 [61%] of the 28 total decisions in this area of opportunity, irrespective of software design limitations). CPOL also displayed eight flags (16% of alerts) that were "false positive" according to mentor recommendations. Overall, CPOL and mentor decisions were the same in 81 of 93 decisions (87%). CPOL agreement with care coordinators was moderate to poor (κ, 0.41; 95% CI, 0.24–0.58), with CPOL flagging decisions more often than care coordinators included the service in the care plan (McNemar's paired proportion test, 95% CI, 9%–30%; P < 0.001).

DISCUSSION

In comparing the aggregate decisions of three mentors with those of the care coordinators in designing a care plan for patients with chronic obstructive pulmonary disease, we found that mentors recommended more services than care coordinators. Care coordinators omitted additional services on 28 occasions when the mentors chose to include them. These 28 decisions represent the opportunity for on-line decision support to improve the alignment of GP and specialist thinking on care planning (beyond simply making it more efficient). Our Care Plan On-Line software flagged 17 of these 28 decisions.

The inconsistency in care planning decisions among the mentors is surprising considering that the mentors worked together on the evidence-based criteria for the services, had identified the services as major issues for patient management, and had jointly designed the initial medical assessment form. That two years elapsed between the authoring of the guidelines and their use in this study probably promoted inconsistency.

Mentor agreement was poor regarding five services. In two of these, inconsistency appears to result from unresolved differences of opinion. For example, one mentor is an expert in bone density issues; such increased preference with increased familiarity is very similar to an observed order of magnitude greater preference for bone densitometry among bone physicians than among respiratory physicians[12] (it is also notable that the Quality of Evidence rating for bone density screening is only III: Box 2). Inconsistency regarding the other three

services appeared to be because of lack of clearly presented clinical information for making a decision. Disagreement on pneumococcal vaccination is particularly illustrative of these problems. The indications for this service are very straight-forward in patients with chronic obstructive pulmonary disease and the mentors all accept the National Health and Medical Research Council (NHMRC) guidelines on this point.[13] However, vagueness in records led to variation in

4: Levels of agreement of care coordinators and CPOL with mentor decisions

	Mentor	Care coordinators			CPOL		
		No	Yes	κ	No	Yes	κ
Counselling on diet/weight/ alcohol/exercise	No	–	–	–	4	0	0.40†
	Yes	–	–		4§	4	
Dietitian	No	9	0	0.16	7	2	0.49†
	Yes	6	1		2	5	
Influenza vaccination	No	1	0	1.00*	0	1	–0.06¶
	Yes	0	18		1	17	
Pneumococcal vaccination	No	0	1	–0.11	1	0	1.00*
	Yes	6	11		0	17	
Bone density screening	No	10	0	0.32†	5	5§	0.45†
	Yes	5	2		0	7	
ECG to look for pulmonary hypertension	No	1	1	0.00	2	0	0.00
	Yes	3	3		1	0	
Smoking cessation	No	16	0	0.62*	16	0	1.00*
	Yes	2	2		0	4	
Physiotherapy	No	4	0	0.28	–	–	–
	Yes	6	4		–	–	
Pharmacy medication management	No	–	–	–	6	0	1.00*
	Yes	–	–		0	9	
Overall	No	41	2	0.49†	41	8	0.72*
	Yes	28‡	41		8	63	

* Relatively good (or perfect) agreement. † Moderate agreement. ‡ Mentors more inclined toward action than care coordinators. §CPOL more inclined to suggest bone density test, but less likely to flag counselling. ¶ κ has a counterintuitive high sensitivity to disagreement in the less common decision direction. CPOL = Care Plan On-Line.

5: Three-way agreement between mentors, CPOL and care coordinators

		Mentor				
		No		Yes		
		CPOL		CPOL		
	Care coordinators	No	Yes	No	Yes	κ
Dietitian	No	7	2	2	4	0.22
	Yes	0	0	0	1	
Influenza vaccination	No	0	1	0	0	0.29
	Yes	0	0	1	17	
Pneumococcal vaccination	No	0	0	0	6	–0.03
	Yes	1	0	0	11	
Bone density screening	No	5	5	0	5	0.19
	Yes	0	0	0	2	
ECG to look for pulmonary hypertension	No	1	0	1	0	–0.29
	Yes	1	0	0	0	
Smoking cessation	No	16	0	0	2	0.76
	Yes	0	0	0	2	
Overall	No	29	8	3	17	0.55
	Yes	2	0	1	33	

Figures in bold italics indicate decisions in which mentors, care coordinators and CPOL agreed. Services where data were not available for all three groups have been omitted from this table. CPOL = Care Plan On-Line.

whether vaccination should be included in the next 12-month care plan. A redesign of the initial medical assessment form to better align with the guideline indicators could improve both specialist consensus and the ability of CPOL to automatically flag services.

The guideline development process for the SA HealthPlus respiratory projects had many of the characteristics necessary for successful implementation, such as involvement of local opinion leading specialists and GPs.[14] Nevertheless, more than 40% of services chosen by the specialists did not appear in the care plans devised by the GPs. To be effective, guidelines need to be integrated into doctors' decision-making processes in daily practice. Dissemination of paper guidelines to GPs is not sufficient for achieving compliance with guidelines in practice.[15] In light of the enormous body of relevant information that GPs might read in an ideal world[16] and the limited available reading time,[17] it is not surprising that guidelines presented as passive reading material have little impact.

Although agreement among the specialists (based on the patient record) was only moderate, the specialist consensus was reasonably well matched by the CPOL algorithms, which indicates that the specialist viewpoint is a coherent target. However, this does not guarantee that it is the best patient management decision.

The GP, in having direct access to the patient, had access to the most complete information. Nearly all discrepancies were in the direction of less service by the GP; responsiveness to patient choice or other practicalities may have contributed to this. An on-line feedback facility would be very helpful to clarify GP motivations for providing fewer services than appear to be indicated by evidence.

The CPOL flags are a moderate to good indicator of aggregate specialist preference in care planning. It would be desirable to improve the alerting capacity of the software through improved indicators on the initial medical assessment form. The observed false-positive alert rate (16%) compares favourably with that for a successful hospital system for preventing adverse drug events (27%).[18] The largest source of CPOL's false-positive errors related to bone

density screening: CPOL presented a "?" flag for the bone-thinning risk factor "female". This naïve interpretation of the guideline was overly sensitive and has been revised to require a higher threshold of risk before presenting an attention flag.

CONCLUSIONS

The view that computers should act as surrogate experts — or as a "Greek Oracle" to be obeyed by the doctor — has long been discredited.[19] A better view is that an intranet-based system allows specialists to communicate their perspective to the GP at the time of decision making. The GP can see the decision specialists would likely take (to include a service in the care plan or not), and, by following a hypertext link, can immediately see an outline of the supporting reasoning and evidence. This leverages the power of Internet/Web systems to disseminate up-to-date information to geographically dispersed communities while avoiding the difficulty of searching for particular facts in a large information space. The technology allows a "mentoring" group to project a specific (perhaps novel) model of care to doctors in the community for their consideration and feedback.

Recent developments — such as the imminent start of a second round of Coordinated Care trials, the *Health Information Network for Australia* report,[20] and the initiatives of the General Practice Computing Group to increase clinical use of computers by GPs — make the time ripe for considering the potential of computer-based aids for dissemination of care-planning guidance. Our study has been limited by surprising levels of specialist disagreement, missing decision data in records, and (in light of these complications) inadequate sample size to arrive at conclusive findings. However, our observations provide strong support that GP decisions with patients are more conservative in chronic respiratory disease care planning than specialist opinion or automated decision support based on review of records. Further investigation of this phenomenon, and the role of computer technology in arriving at true best practice in care planning, is warranted.

REFERENCES

1. Leape LL. Practice guidelines and standards: an overview. *QRB Qual Rev Bull* 1990; 16: 42-49.
2. Wall E. Practice guidelines: promise or panacea? *J Fam Pract* 1993; 37: 17-29.
3. Lobach D. A model for adapting clinical guidelines for electronic implementation in primary care. *J Am Med Inform Assoc* 1995; Symposium Suppl: 581-585.
4. Grimshaw J, Russell I. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993; 324: 1317-1322.
5. Shiffman R. Representation of clinical practice guidelines in conventional and augmented decision tables. *J Am Med Inform Assoc* 1997; 4: 382-393.
6. Zielstorff R. Online practice guidelines. *J Am Med Inform Assoc* 1998; 5: 227-236.
7. Commonwealth Department of Health and Aged Care. HealthPlus Coordination Care Trial (SA). Updated 11 Jan 2001. <http://www.health.gov.au/hscd/primcare/first/healthplus.htm>. Accessed 21 August 2001.
8. Warren J, Beliakov G, Noone J, Frankel H. Chronic disease coordinated care planning: flexible, task-centered decision support. *Top Health Inform Manage* 1999; 20: 52-68.
9. Eccles M, Clapp Z, Grimshaw J, et al. North of England evidence based guideline development project: methods of guideline development. *BMJ* 1996; 312: 760-762.
10. Morton A, Dobson A. Assessing agreement. *Med J Aust* 1989; 150: 384-387.
11. Fleiss J. Statistical methods for rates and proportions. New York: Wiley, 1981; 212-34.
12. Connolly C, Francis R, Murth N, et al. Steroid-induced bone disease. *J R Coll Physicians Lond* 1992; 26: 278-280.
13. Watson C, editor. The Australian Immunisation Handbook, 6th ed. Canberra: Australian Government Publishing Service, 1997.
14. Lomas J, Enkin M, Anderson G, et al. Opinion leaders vs audit and feedback to implement practice guidelines. *JAMA* 1991; 265: 2202-2207.
15. Lam S. Implementation and evaluation of practice guidelines. *J Am Med Inform Assoc* 1994; Symposium suppl: 253-263.
16. Davidoff F, Haynes B, Sackett D, Smith R. Evidence-based medicine; a new journal to help doctors identify the information they need. *BMJ* 1995; 310: 1085-1086.
17. Sackett D, Rosenberg W, Gray J, et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312: 71-72.
18. Raschke R, Gollihare B, Wunderlich T, et al. A computer alert system to prevent injury from adverse drug events. *JAMA* 1998; 280: 1317-1320.
19. Miller R, Maserie F. The demise of the 'Greek Oracle' model for medical diagnostic systems. *Methods Inf Med* 1990; 29: 1-2.
20. Commonwealth Department of Health and Aged Care. A health information network for Australia. Report of the National Electronic Health Records Task Force. Canberra: Commonwealth of Australia, 2000. Available at: <http://www.health.gov.au/healthonline/ehr_rep.htm>.