

Articles

Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations

Jeremy M Grimshaw, Ian T Russell

Summary

Although interest in clinical guidelines has never been greater, uncertainty persists about whether they are effective. The debate has been hampered by the lack of a rigorous overview.

We have identified 59 published evaluations of clinical guidelines that met defined criteria for scientific rigour; 24 investigated guidelines for specific clinical conditions, 27 studied preventive care, and 8 looked at guidelines for prescribing or for support services. All but 4 of these studies detected significant improvements in the process of care after the introduction of guidelines and all but 2 of the 11 studies that assessed the outcome of care reported significant improvements.

We conclude that explicit guidelines do improve clinical practice, when introduced in the context of rigorous evaluations. However, the size of the improvements in performance varied considerably.

Lancet 1993; **342**: 1317–22

Introduction

Interest in clinical guidelines has never been greater.¹ The hope is that guidelines will reduce inappropriate practice and so improve efficiency² but there is concern that they might be insensitive to the needs of individual patients and uncertainty about whether they do change practice. On one thing there is general agreement: the widespread introduction of clinical guidelines will cost money. These issues—uncertainty about effectiveness, concern about side-effects, and worries about costs—are seen with new health care technologies too and both types of development must be critically evaluated.

Another useful analogy between guidelines and technologies (including drugs) is that if effectiveness is not obvious from the first few randomised trials a systematic review or meta-analysis may be required for a firm conclusion. Examples of technologies successfully evaluated in this way include beta blockade after myocardial infarction, prophylactic antibiotics in caesarean

section, perioperative parenteral nutrition, and spinal manipulation for back and neck pain. However, clinical guidelines have yet to be subjected to this sort of scrutiny. Reviews of continuing medical education,^{4–6} computer-aided quality assurance,^{5,7} drug-oriented quality assurance,⁸ and feedback on clinical performance^{9,10} have all included *en passant* some of the published evaluations of clinical guidelines. Two reviews that did focus on guidelines were restricted to consensus statements¹¹ or to guidelines in general practice.¹² We therefore searched for published evaluations of clinical guidelines that met defined criteria for scientific rigour.

Methods

Systematic review of literature

Our definition of clinical guidelines is “systematically developed statements to assist practitioner decisions about appropriate health care for specific clinical circumstances”.¹³ Thus we excluded sets of criteria for the appropriateness of individual items of care that are not integrated into coherent guidelines, educational programmes that did not generate guidelines, quality assurance not associated with guidelines, and feedback on clinical performance not related to specific guidelines. We also excluded patient-oriented strategies for improving compliance with guidelines; and in studies of both doctor and patient oriented strategies we concentrated on the doctors.

We searched: (1) computerised databases (Medline, the UK Department of Health’s DHSS-DATA, Elsevier’s Embase, the US National Library of Medicine’s Grateful Med, and the Dutch EAGLE Technical Committee’s “grey literature” system SIGLE); (2) published bibliographies of related topics;^{4–12} (3) citations in articles reviewed; and (4) references provided by colleagues. For computer searches we used the following terms: algorithm, consensus statement, controlled trial, health planning guidelines, health services research, medical audit, practice guidelines, process and outcome evaluation, quality assurance, quality of health care, and reference standard. We also searched for key authors’ names. We limited our review to publications that had appeared or were known to us by the end of 1992. We reviewed in depth only those investigations of clinical guidelines designed for medical staff that evaluated the introduction of guidelines in terms of the process of medical care or the outcome for patients, collected enough data for meaningful statistical analysis, and, most important of all, had a rigorous study design.

Criteria for study designs

Meta-analyses typically review only randomised trials. We chose to include evidence from other robust designs because the large improvement in the process of care that many of them show is less likely than smaller improvements to be entirely attributable to bias and because the simple randomised trial cannot be regarded as the gold standard in behavioural research.¹⁴ Randomised trials are the most robust method of assessing most health care innovations, but in evaluating guidelines there is a danger that treatment offered to the control patients will be contaminated by doctors’ knowledge of

Department of General Practice, University of Aberdeen, Foresterhill Health Centre, Westburn Road, Aberdeen AB9 2ZD, UK
(J M Grimshaw MRCGP); Health Services Research Unit, University of Aberdeen (I T Russell PhD)

Correspondence to: Dr Jeremy M Grimshaw

Author(s) (year)*	Disease	Intervention (dissemination or implementation)	Design	Effects on process of care
McDonald (1976)	Various	External; computer-generated reminders	Randomised by patient	Compliance with investigation guidelines increased by 25% (11% to 36%) and therapeutic guidelines by 15% (13% to 28%) Compliance increased by 29% (22% to 51%)
McDonald (1976)	Various, managed or caused by drugs	External; computer-generated reminders	Crossover, randomised by doctor	
Barnett et al (1978)	Streptococcal sore throat	Intermediate; patient specific feedback	Time series, switchback design	Patients untreated at 10 days fell by 10% (12% to 2%)
Sanazaro and Worth ²⁸ (1978)	Various	National guidelines placed in patients' records	Randomised by hospital	Compliance with treatment guidelines increased by 2% (87% to 89%) (significant) but no significant change in compliance with documentation guidelines Compliance increased by 37% (45% to 82%) Average non-compliance fell from 4.3 to 2.8 items
Hopkins et al ²⁴ (1980)	Hypotensive shock	External; personal instruction	Randomised by team	
Linn ²⁵ (1980)	Burns	National; seminar	Randomised by hospital	
McDonald et al (1980)	Various	External; computer-generated reminders with or without bibliographical citations	Crossover, randomised by doctor	Compliance increased by 18% (20% to 38%) but no significant increase following bibliographical citations
Barnett et al ²⁶ (1983)	Hypertension	External; patient-specific feedback	Randomised by patient	Follow-up increased by 52% (46% to 98%) and blood pressure recording increased by 18% (52% to 70%) Compliance increased by 13% (37% to 50%)
Thomas et al ³⁰ (1983)	Diabetes	External; computer-generated reminders	Randomised by patient	
Sommers et al (1984)	Unexplained anaemia	(1) internal or external; aggregated feedback (2) same; computer-generated reminders	(1) randomised by doctor (2) uncontrolled before and after	(1) compliance increased in group receiving external guidelines by 14% (37% to 51%) (2) compliance increased by 30% (36% to 66%)
Norton & Dempsey ¹⁵ (1985)	Cystitis and vaginitis	Internal; aggregated feedback	Latin square, randomised by doctor	Compliance increased for cystitis by 28% (38% to 66%) and for vaginitis by 9% (22% to 31%)
Palmer et al (1985)	Various	Intermediate; aggregated feedback	Balanced incomplete block, randomised by practice	Compliance increased by 10% for gastroenteritis care (20% to 30%) and by 7% for well child care (22% to 29%); no significant increase for remaining 4 conditions Compliance increased by 32% or 22%
Putnam & Curry ¹⁶ (1985)	Various	Internal or external group; aggregated feedback	Randomised by doctor	
Brownbridge et al ¹⁴ (1986)	Hypertension	Intermediate; computerised protocol	Before and after, controlled by consultations at branch surgery	Compliance increased from 2.3 to 8.4 items
McAlister et al ²³ (1986)	Hypertension	Provincial; patient-specific feedback	Randomised by doctor	Study group treated higher proportion of patients and followed them for longer; differences not significant
Wirtshafter et al (1986)	Neonatal respiratory distress syndrome	External; lectures and embedded within new records	Randomised by hospital	Compliance increased by 8% (50% to 58%) but no significant changes in hospitals receiving only lectures
Kosecoff et al ²² (1987)	Breast cancer, caesarean section, coronary artery bypass grafting	Consensus development conference; publication in journals	Interrupted (uncontrolled) time series	No significant change in rate of increase in compliance with 11 recommendations from four conferences
Lomas et al ²⁰ (1989)	Caesarean section	Consensus development conference; mail to targeted doctors	Interrupted (uncontrolled) time series	Caesarean section rate (which had been increasing by 0.3% per annum) began to decrease by 0.5% per annum; Caesarean rate after previous section (which had been increasing by 1.5%) began to decrease by 2.2%
Lomas et al (1991)	Caesarean section	Consensus development conference; opinion leader or aggregated feedback	Randomised by hospital	Compliance increased in hospitals with opinion leader (eg, elective sections decreased by 15%, from 69% to 54%); no significant change in hospitals receiving feedback
Durand-Zaleski et al (1992)	Hypovolaemia	Consensus development conference; package of local initiatives	Interrupted (uncontrolled) time series	Use of albumin (which had increased by more than 100% over 2 years preceding conference) decreased by 40% over the following year and remained constant thereafter
Margolis et al (1992)	Various, paediatric	External; computerised protocols	Partly balanced incomplete block, randomised by doctor	Compliance increased for data collection by 57% (42% to 99%) and for management by 20% (42% to 62%)
North of England Study ²⁸ (1992)	Various, paediatric	Internal; mail to other groups	Latin square, randomised by practice	Developing (but not receiving) guidelines improved compliance for all five conditions, prescribing by 8% on average and follow-up by 6% on average
Sherman et al (1992)	Prostatic carcinoma	Consensus development conference; publication in journals	Interrupted (uncontrolled) time series	No significant change in rate of increase in compliance with recommendations relating to prostatectomy and radiotherapy
Emslie et al (1993)	Infertility	Intermediate; embedded within new records	Randomised by practice	Compliance increased by between 6% (80% to 86%) for asking about dyspareunia and 39% (25% to 64%) for investigation of rubella status

*A list of references is available from *The Lancet* or the authors.**Table 1: Studies of effects of guidelines on process of clinical care**

the guidelines, leading to underestimates of the true effects of guidelines. Studies where doctors (or hospitals) are randomised are at risk of a different bias: those randomised to the guidelines group may be subject to a greater Hawthorne effect (the beneficial effect on performance of taking part in research) than controls, with the result that the evaluation may overestimate the true effects of guidelines. One way of countering this bias is a crossover trial.

Balanced incomplete block designs provide an even more powerful way of overcoming many of the problems of simple randomised studies—by targeting different doctors for different activities selected at random according to a robust experimental plan. For example, Norton and Dempsey¹⁵ used a 2×2 latin square to investigate guidelines for two common gynaecological complaints in a Canadian family practice. Study doctors were

Author(s) (year)*	Area	Intervention (dissemination or implementation)	Design	Effects on process of care
Morgan et al (1978)	Antenatal care	National; computer-generated reminders	Randomised by patient	Compliance increased in experimental patients by 15% (83% to 98%) within six mo and in control patients by 11% (83% to 94%) within 12 mo
Cohen DI et al ¹⁷ (1982) Rodney et al (1983)	Preventive care Immunisations	External; embedded in patients' records External; redesigned record	Randomised by team Before and after, controlled by influenza immunisations (untargeted)	Compliance increased by 32% (4% to 36%) Targeted immunisations increased by 15% (3% to 18%); untargeted immunisations remained at 5%
Thompson et al ²⁸ (1983)	Chest X-rays and multi-channel blood tests in 'routine' physical examinations	Intermediate; extended educational programme	Before and after, controlled by national data	Non-indicated X-rays decreased by 24% (30% to 6%) while national usage remained constant; non-indicated blood tests decreased by 21% (31% to 10%)
McDonald et al ²⁷ (1984) Winickoff et al (1984)	9 preventive tasks, 6 laboratory tests Screening for colorectal cancer	External; computer-generated reminders (1) internal; aggregated feedback (2) doctor specific feedback	Randomised by doctor (1) uncontrolled before-and-after (2) crossover, randomised by doctor	Average compliance increased by 20% (29% to 49%) (1) no significant increase in compliance (2) compliance increased by 10% (72% to 82%)
Cohen SJ et al ²¹ (1985)	13 preventive tasks	External; financial incentive (credit at bookshop)	Latin square, randomised by doctor	Doctors' knowledge scores increased by 9% (53% to 62%)
McDowell et al (1986)	Influenza vaccination	National; computer-generated reminder	Randomised by patient	Compliance increased by 13% (10% to 23%)
Prislin et al (1986)	Stool occult blood testing and breast examination	External; flowsheet in patients' records	Randomised by patient	Compliance increased for breast exam by 32% (40% to 72%) and for occult blood by 15% (39% to 54%); latter not significant
Tierney et al (1986)	11 preventive tasks	External; computer-generated reminder or patient specific feedback	Balanced incomplete block design, randomised by doctor	Reminders increased compliance by 15% (15% to 30%); feedback increased compliance by 7% (15% to 22%)
Cheney and Ramsdell (1987) Cohen SJ et al (1987)	12 preventive tasks Smoking	National; checklist in patients' records National; 2 types of patient-specific reminder	Randomised by doctor 2 x 2 trial, randomised by doctor	Compliance increased by 17% (39% to 56%)
Robie (1988)	3 preventive tasks	National; patient specific reminders	Before and after; controlled by second group of doctors	Compliance increased by 32% (23% to 55%) with one reminder and 56% (to 79%) with two reminders (as reported by patients)
Schreiner et al (1988)	4 preventive tasks	National; patient-specific reminders	Switchback; controlled by second group of doctors	Compliance increased by 28% (22% to 50%) for cervical smears; no significant increase in compliance for breast or rectal examinations
Wilson et al (1988)	Smoking	External; recruiting patient through receptionists	Randomised by practice	Compliance increased by 8% (32% to 40%)
Becker et al (1989) Chambers et al (1989)	9 preventive tasks Mammography	National; patient-specific reminder National; computer-generated reminders	Randomised by patient Randomised by patient	Compliance increased by 64% (12% to 76%) as reported by patients Compliance increased by 5% (8% to 13%)
Cummings et al (1989)	Smoking	External; patient-specific reminders	Randomised by practice	Compliance increased for study patients by 13% (14% to 27%) and for controls by 7% (14% to 21%)
McDowell et al (1989)	Blood pressure screening	National; computer-generated reminders	Randomised by patient	Patients reported discussion of smoking increased by 5% (45% to 50%) and compliance with guidelines increased by 12% (10% to 22%) when smoking discussed
McDowell et al (1989)	Cervical screening	National; computer-generated reminders	Randomised by patient	Compliance increased by 10% (21% to 31%)
McPhee et al (1989)	7 preventive tasks	National; computer-generated reminders or aggregated feedback	Randomised by doctor	No significant increase in compliance
McPhee et al (1991)	11 preventive tasks	National; computer-generated reminders	Randomised by doctor	No significant increase in compliance for cervical smears; reminders increased average compliance for other six tasks by 22% (49% to 71%); feedback increased this compliance by 14% (49% to 63%)
Rosser et al (1991)	Smoking	National; computer-generated reminder*	Randomised by patient	No significant increase in compliance for mammography or sigmoidoscopy; compliance for other nine tasks increased by 14% (39% to 53%)
Cowan et al (1992) Headrick et al (1992)	7 preventive tasks Cholesterol	National; placed in patients' records National; placed in patients' records or implemented by computer-generated reminders	Randomised by doctor Randomised by doctor	Compliance increased by 26% (12% to 38%)
Lilford et al (1992)	Antenatal care	Intermediate; new paper record or computerised questionnaire	Randomised by patient	Compliance increased by 5% (2% to 7%)
Rosser et al (1992)	Tetanus vaccination	National; computer-generated reminder	Randomised by patient	Guidelines in records increased by 8% (43% to 51%) and computer-generated reminders increased compliance by 11% (36% to 47%) Compliance increased by 8% for both

*A list of references is available from *The Lancet* or the authors.

Table 2: Studies of effects of guidelines on process of preventive care

randomly allocated between two groups, of which one developed guidelines for the care of cystitis, while the other developed guidelines for vaginitis; as each group acted as a blind control for the other, doctors in both groups enjoyed the same Hawthorne effect.

Controlled before-and-after studies aim to identify a non-randomised control group that will experience trends and changes similar to those of the study population, thus overcoming the difficulties of uncontrolled before-and-after studies and reducing

the risk of Hawthorne effects. The control may be either an external group of doctors similar to the study group or activities performed by study doctors but not targeted by guidelines. We included before-and-after studies with controls but excluded uncontrolled ones because any secular trend or sudden change makes it difficult to attribute observed changes to the intervention. Before-and-after studies can also be strengthened by time-series analysis or by observing patient care before, during, and after a well-defined intervention period (a switchback design).

Author(s) (year)*	Service	Intervention (dissemination or implementation)	Design	Effects on process of care
Brook et al ^{18,19} (1976)	Injectable antibiotic prescribing for respiratory conditions	Intermediate; financial incentive (payment for compliance)	Time series controlled by untargeted prescribing	Prescribing of injections decreased by 60% while all other prescribing increased by 3%; compliance with criteria for antibiotic prescribing increased in six respiratory conditions by between 6% and 39%
Fowkes et al (1984)	Skull X-rays for head injury	National; new record	Uncontrolled time series	Use of X-rays decreased by 51% (65/1000 to 32/1000)
Fowkes et al (1986)	Preoperative chest X-rays	National; variety of local initiatives	Before and after; in four hospitals controlled by a fifth	X-ray use reduced by between 8% and 16%
Landgren et al (1988)	Antibiotic prophylaxis in surgery	Intermediate; marketing campaign and aggregated feedback	Crossover; randomised by hospital	Compliance increased: duration of prophylaxis by 21% (39% to 60%); timing of prophylaxis by 18% (48 to 66%); latter not significant
Bareford and Haying (1990)	Haematological tests	External; aggregated feedback and cancellation of inappropriate expensive tests	Uncontrolled time series	Haematology requests fell by 20%, departmental falls ranging from zero to 38%
Clarke and Adams (1990)	Skull X-rays for head injury	Intermediate; posters and lectures	Uncontrolled time series	Use of X-rays decreased by 27% (94/1000 to 68/1000)
De Vos Meiring and Wells (1990)	Radiological investigations	External; mail to targeted doctors	Before and after; controlled by five untargeted radiological investigations	Targeted investigations decreased by 28% while untargeted investigations decreased by 2%
Gama et al (1992)	Cardiac enzyme tests	Intermediate; seminar	Time series	Lactate dehydrogenase (not recommended) virtually eliminated; creatine kinase (recommended for 2 days rather than 3) unchanged

*A list of references is available from *The Lancet* or the authors.

Table 3: **Studies of effects of guidelines on prescribing and use of support services**

For these reasons we limited our review to evaluations that used one or more of five study designs: balanced incomplete block, randomised crossover, simple randomised, controlled before and after (including switchback designs), and interrupted time series.

Results

59 evaluations of guidelines met these criteria (24 on specific clinical conditions, 27 on prevention, and 8 on prescribing or on laboratory or radiological services). Tables 1–3 summarise the studies and their effects on process of care (did they lead to changes in what doctors do?). Table 4 summarises effects on outcome (did the guidelines lead to improvements in patients' health?) for the 11 studies that investigated this. To facilitate comparisons the effects of the guidelines on compliance are expressed as absolute percentages where possible.

Author(s) (year)*	Effects on outcome of care
Sanzaro and Worth ²⁹ (1978)	No significant change
Hopkins et al ²⁴ (1980)	Patients requiring ventilation reduced by 19% (33% to 14%)
Linn ²⁵ (1980)	Early complications in patients admitted to study hospitals reduced by 15% (45% to 30%) and patient non-compliance after 2 wks in study hospitals reduced by 5% (10% to 5%)
Barnett et al ²⁹ (1983)	After 2 yrs patients with diastolic BP < 100 mm Hg or on treatment increased by 18% (52% to 70%)
Thomas et al ³⁰ (1983)	No significant change
McDonald et al ²⁷ (1984)	Patients eligible for pneumococcal or influenza vaccination suffered fewer winter hospitalisations and emergency room visits during years with influenza epidemics
McAster et al ²³ (1987)	In newly detected hypertensives number of days per year with diastolic BP < 90 mm Hg increased by 19% (from 255 to 323)
Wilson et al ²⁸ (1988)	Patients who reported not smoking after one year increased by 4.4% (4.4 to 8.8%)
Cohen SJ et al ^{29,30} (1989)	Patients not smoking after 1 yr increased by 8.4% (2.7% to 11.1%)
Cummings et al ³¹ (1989)	Patients not smoking after 9 mo increased by 1.1% (1.5% to 2.6%) (not significant); motivated patients not smoking after 9 mo increased by 2.2% (1.7% to 3.9%)
North of England Study ²⁸ (1992)	Developing (but not receiving) internal guidelines improved patient outcomes for only one of the five study conditions (recurrent wheezy chest) including: patient compliance increased by 14% (79% to 91%) and patient breathlessness decreased by 40% (4.2 to 1.7 days/mo)

*A list of references is available from *The Lancet* or the authors.

Table 4: **Studies of effects of guidelines on patient outcome**

Effects on process of care

All but 4 of the 59 studies detected significant change in the process of care in the direction proposed by the guidelines. The size of this improvement varied considerably. For example, Putnam and Curry¹⁶ evaluated clinical guidelines on common medical complaints seen in sixteen Canadian family practices; they reported a 32% increase in compliance when family physicians had been involved in developing guidelines but only a 22% increase when the guidelines had been developed by others (table 1). Cohen et al¹⁷ evaluated the effects of attaching guidelines for preventive care to patients' notes; compliance increased by 32% (table 2). Brook and colleagues^{18,19} studied guidelines restricting the use of antibiotic injections for respiratory infections when reinforced by a decision of the funding body (Medicaid) that claims not meeting the guidelines would not be reimbursed; the prescription of antibiotic injections fell by 60% while that of other prescriptions hardly changed (table 3). Lomas et al²⁰—in a Canadian study of caesarean section that is often cited as evidence of the ineffectiveness of guidelines—did detect a small but significant change in caesarean section rates in the direction of the guidelines (table 1). Cohen et al²¹ examined the effects of "reading" (ie, studying scientific papers) on residents' compliance with preventive care guidelines; they found only a "modest" increase in compliance (table 2).

In one of the 4 studies that detected no improvement in the process of care, Kosecoff et al²² evaluated the effect of four American consensus conferences by surveying ten randomly selected hospitals in the State of Washington.

Probability of being effective	Development strategy	Dissemination strategy	Implementation strategy
High	Internal	Specific educational intervention	Patient-specific reminder at time of consultation
Above average	Intermediate	Continuing education	Patient-specific feedback
Below average	External, local	Mailing targeted groups	General feedback
Low	External, national	Publication in journal	General reminder

Table 5: **Classification of clinical guidelines**

Time-series analysis showed that, although compliance with the guidelines continued to increase after the conferences, the rate of increase fell, suggesting that the conferences had no effect on practice (table 2). McAlister et al²³ evaluated the effects of computer feedback on hypertension management in Canadian family practice; although they found no significant improvement in the process of care (table 1), they did report a significant improvement in the control of blood pressure (an "intermediate outcome") in intervention practices (table 4).

Effects on outcome of care

All but 2 of the 11 studies of patient outcome found some significant improvement (table 4). Hopkins et al²⁴ evaluated guidelines for the management of hypotensive shock in a US emergency room; care complied with guidelines in 82% of patients treated by residents in the intervention group compared with 45% treated by the control group (table 1) and of patients treated by the intervention group only 14% required ventilation compared with 33% treated by the control group (table 4). Linn²⁵ evaluated guidelines for the management of burns in twenty US emergency rooms; improved compliance with guidelines in intervention hospitals (table 1), was followed by improved patient compliance with treatment and a reduction in early complications (table 4). Barnett et al²⁶ evaluated the effects of computer-generated reminders on the management of poorly followed up hypertensive patients in a US health centre; more patients with reminders were followed up (98% vs 46%) and had their blood pressures recorded (70% vs 52%) and had a better outcome at two years, defined as diastolic less than 100 mm Hg or on treatment (also 70% vs 52%). McDonald et al²⁷ examined the effects within a US internal medicine clinic of computer-generated reminders for fifteen activities (nine related to prevention and six to laboratory testing). Compliance was 20% greater in residents randomly allocated to receive reminders and patients eligible for influenza vaccination were less likely to visit the emergency room or to be admitted to hospital during years when influenza epidemics occurred. Three US studies²⁸⁻³¹ that evaluated guidelines for helping smokers to give up all achieved substantial improvements in the process of care and significant reductions in the proportion of patients smoking after 9-12 months (tables 2 and 4). In the North of England Study of Standards and Performance in General Practice,³² general practitioners set standards for the care of five common paediatric conditions. When general practitioners developed internal guidelines (but not when they received external guidelines), the process of care improved for all five and the outcome for one of them drug compliance for recurrent wheezy chest improved by 14% and parents reported a decrease in breathlessness from 4.2 to 1.7 days per month (table 4).

In contrast, Sanazaro and Worth³³ reported only a 2% increase in compliance with treatment guidelines for seven study conditions and failed to detect any change in outcome following dissemination of guidelines; however, they did note that failure to comply with the guidelines was associated with a poorer outcome for patients with pneumonia or acute myocardial infarction. In evaluating the effect of computer generated-reminders on diabetic care, Thomas and colleagues,³⁴ showed significantly improved compliance with guidelines; although patients with reminders experienced fewer hospital admissions than controls, this difference was not significant.

Discussion

Of 59 papers that evaluated clinical guidelines through rigorous research designs all but 4 detected significant improvements in the process of care following the introduction of guidelines. Although the clinical significance of a few of these papers is questionable, others showed substantial improvement in quality of care and 9 of the 11 papers that assessed the outcome of care reported significant improvements. The conclusion is that explicit guidelines do improve clinical practice, in the context of rigorous evaluations.

The successful introduction of clinical guidelines is dependent on many factors including the clinical context and the methods of developing, disseminating, and implementing those guidelines.¹⁴ Different methods will be appropriate in different contexts. Studies that report large improvements in clinical care suggest the potential of guidelines when development, dissemination, and implementation are all appropriate. Studies that report small improvements or none may reflect failure at any stage during the introduction or evaluation of the guidelines. We therefore suggested a classification that highlights the more effective strategies (table 5).¹⁴ Only if appropriate strategies are selected at each stage will clinical guidelines achieve the full potential that this review has demonstrated.

We thank the Wellcome Trust for funding JMG and the Chief Scientist Office of the Scottish Home and Health Department for funding the Health Services Research Unit (ITR). We thank Sheila Wallace for help with the literature review and Mark Garton, Simon Naji and Lewis Ritchie for comments on an earlier draft.

References

- Haines A, Feder G. Guidance on guidelines. *BMJ* 1992; **305**: 785-86.
- Irvine D. Managing quality in general practice. London: King's Fund, 1991.
- Editorial. Guidelines for doctors in the new world. *Lancet* 1992; **339**: 1197-98.
- Haynes RB, Davis DA, McKibbon A, Tugwell P. A critical appraisal of the efficacy of continuing medical education. *JAMA* 1984; **251**: 61-64.
- Lomas J, Haynes RB. A taxonomy and critical review of tested strategies for the application of clinical practice recommendations. *Am J Prev Med* 1987; **4**: 77-94.
- Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME: a review of 50 randomised controlled trials. *JAMA* 1992; **268**: 1111-17.
- Haynes RB, Walker CJ. Computer-aided quality assurance: a critical appraisal. *Arch Intern Med* 1987; **147**: 1297-301.
- Soumerai SB, McLaughlin TJ, Avorn J. Improving drug prescribing in primary care: a critical analysis of the experimental literature. *Millbank Q* 1989; **67**: 268-317.
- Mugford M, Banfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. *BMJ* 1991; **303**: 398-402.
- Buntinx F, Winkens R, Grol R, Knottnerus JA. Influencing diagnostic and preventive performance in ambulatory care by feedback and reminders: a review. *Fam Pract* 1993; **10**: 219-28.
- Lomas J. Words without action? the production, dissemination and impact of consensus recommendations. *Annu Rev Public Health* 1991; **12**: 41-65.
- Grol R. Implementing guidelines in general practice care. *Qual Health Care* 1992; **1**: 184-91.
- Field MJ, Lohr KN. Clinical practice guidelines: direction of a new program. Washington DC: National Academy Press, 1990.
- Russell IT, Grimshaw JM. The effectiveness of referral guidelines: a review of the methods and findings of published evaluations. In Roland M, Coulter A, eds. Hospital referrals. Oxford: Oxford University Press, 1992: 179-211.
- Norton PG, Dempsey LJ. Self-audit: its effect on quality of care. *J Fam Pract* 1985; **21**: 289-91.
- Putnam RW, Curry L. Impact of patient care appraisal on physician behaviour in the office setting. *Can Med Assoc J* 1985; **132**: 1025-29.
- Cohen DI, Littenberg B, Wetzel C, Neuhauser DB. Improving

- physician compliance with preventive medicine guidelines. *Med Care* 1982; **20**: 1040-45.
- 18 Brook RH, Williams KN. Effect of medical care review on the use of injections: a study of the New Mexico Experimental Care Review Organisation. *Ann Intern Med* 1976; **85**: 509-15.
 - 19 Lohr KN, Brook RH. Quality of care in episodes of respiratory illness among Medicaid patients in New Mexico. *Ann Intern Med* 1980; **92**: 99-106.
 - 20 Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *N Engl J Med* 1989; **321**: 1306-11.
 - 21 Cohen SJ, Weinberger M, Hui SL, Tierney WM, McDonald CJ. The impact of reading on physicians non adherence to recommended standards of medical care. *Soc Sci Med* 1985; **21**: 909-14.
 - 22 Kosecoff J, Kanouse DE, Rogers WH, McCloskey L, Winslow CM, Brook RH. Effects of the National Institutes of Health consensus development program on physician practice. *JAMA* 1987; **258**: 2708-13.
 - 23 McAlister NH, Covvey HD, Tong C, Lee A, Wigle ED. Randomised controlled trial of computer assisted management of hypertension in primary care. *BMJ* 1986; **293**: 670-74.
 - 24 Hopkins JA, Shoemaker WC, Greenfield S, Chang PC, McAuliffe T, Sproat RW. Treatment of surgical emergencies with and without an algorithm. *Arch Surg* 1980; **115**: 745-50.
 - 25 Linn BS. Continuing medical education: impact on emergency room burn care. *JAMA* 1980; **244**: 565-70.
 - 26 Barnett GO, Winickoff RN, Morgan MM, Zielstorff RD. A computer-based monitoring system for follow-up of elevated blood pressure. *Med Care* 1983; **21**: 400-09.
 - 27 McDonald CJ, Hui SL, Smith DM, et al. Reminders to physician from an introspective computer medical record: a two year randomised trial. *Ann Intern Med* 1984; **100**: 130-38.
 - 28 Wilson DM, Taylor DW, Gilbert R, et al. A randomised trial of a family physician intervention for smoking cessation. *JAMA* 1988; **260**: 1570-74.
 - 29 Cohen SJ, Christen AG, Katz BP, et al. Counselling medical and dental patients about cigarette smoking: the impact of nicotine gum and chart reminders. *Am J Public Health* 1987; **77**: 313-16.
 - 30 Cohen SJ, Stookey GK, Katz BP, Drook CA, Smith DM. Encouraging primary care physicians to help smokers quit: a randomised trial. *Ann Intern Med* 1989; **110**: 686-52.
 - 31 Cummings SR, Coates TJ, Richard RJ, et al. Training physicians in counselling about smoking cessation: a randomised trial of the "Quit for Life" program. *Ann Intern Med* 1989; **110**: 640-47.
 - 32 North of England Study of Standards and Performance in General Practice. Medical audit in general practice: effects on doctors' clinical behaviour and the health of patients with common childhood conditions. *BMJ* 1992; **304**: 1480-88.
 - 33 Sanazaro PJ, Worth RM. Concurrent quality assurance in hospital care. *N Engl J Med* 1978; **298**: 1171-77.
 - 34 Thomas JC, Moore A, Qualls PE. The effect on cost of medical care for patients treated with an automated audit system. *J Med Syst* 1983; **7**: 307-13.

Angioscopy in variant angina: coronary artery spasm and intimal injury

Hirokuni Etsuda, Kyoichi Mizuno, Koh Arakawa, Kimio Satomura, Toshio Shibuya, Kazushige Isojima

Summary

Studies in pigs and dogs show that intimal injury is related to coronary artery spasm; it is not known whether intimal injury is related to coronary artery spasm in human beings.

We examined intima at the site of coronary artery spasm by percutaneous transluminal coronary angioscopy in 10 of 13 consecutive patients with variant angina. Coronary artery spasms occurred spontaneously or were induced by intracoronary acetylcholine (10-100 µg). Angioscopy showed intimal injuries (haemorrhage, flap, thrombus, or ulcer) in 4 of the 10.

We suggest that intimal injury is related to coronary artery spasm in human beings.

Lancet 1993; **342**: 1322-24

Introduction

Attacks in patients with variant angina are often the result of local coronary hyper-reactivity to coronary constrictor stimuli.¹ However, neither a specific agonist inducing coronary spasm nor specific antagonist preventing it have been shown.² It was reported recently that acetylcholine causes vasoconstriction at the site of endothelial injury because acetylcholine-mediated endothelium-derived relaxing factor, NO, was decreased or absent.³ Coronary artery spasm with regional myocardial ischemia has been

shown in pigs after the endothelium was damaged;⁴ scanning and transmission electron microscopy studies showed endothelial injury at the site of coronary artery spasm in dogs;⁵ but this has not been shown in human beings with variant angina.

Percutaneous transluminal coronary angioscopy has been successfully used to investigate coronary intimal macromorphology.⁶ Angioscopy is more sensitive than arteriography for observing intra-coronary structure. This study examines intimal coronary artery morphology at the site of spasm in patients with variant angina.

Patients and methods

Patients

Percutaneous transluminal coronary angioscopy was attempted in 13 consecutive patients with variant angina. Written informed consent was obtained from each patient before admission to the study; complications and their treatments were explained; and published studies^{7,8} were given to them. The protocol was approved by The National Defense Medical College Institutional Review Board. 3 patients were excluded from the study because of inadequate angioscopy pictures. Baseline characteristics and catheterisation data on the 10 patients were: male 9, age 56 (± 6.5); site of angioscopy was left anterior descending coronary artery (6), circumflex coronary artery (1), right coronary artery (3); % diameter of fixed stenosis at the site of angioscopy 29% (± 17); previous medication was calcium blocker (10), nitrate (10), aspirin (9), β blocker (0). Most patients had early-morning rest angina. The chest pain had not progressed from the time of onset, and the symptoms resolved spontaneously or with sublingual nitroglycerine. No patients were on heparin for the treatment of chest pain. The diagnosis of variant angina was based on clinical characteristics, electrocardiographic changes (transient ST

First Department of Internal Medicine, National Defense Medical College, 3-2 Namiki Tokorozawa 359, Japan (H Etsuda MD, K Mizuno MD, K Arakawa MD, K Satomura MD, T Shibuya MD, K Isojima MD)

Correspondence to: Dr Kyoichi Mizuno