- 20 Woo J, Lam CWK, Kay R, Teoh R, Wong HY, Prall WY, *et al*. Hypertension, lipoprotein(a) and apolipoprotein A-I as risk factors for stroke in the Chinese. *Stroke* 1991;22:203-8.
- 21 Paterson JC, Dyer L, Armstrong EC. Serum cholesterol level in human atherosclerosis. *Can Med Assoc J* 1960;82:6-11.
- 22 Smith GD, Shipley MJ, Marmot MG, Rose G. Plasma cholesterol concentration and mortality. *JAMA* 1992;267:70-6.
- 23 Stokes J III, Kannel WB, Wolf PA, Cupples A, D'Agostino RB. The relative importance of selected risk factors for various manifestations of cardiovascular disease among men and women from 35 to 64 years old: 30 years follow-up in the Framingham study. *Circulation* 1987;75(suppl V):65-73.
- 24 Stemmermann GN, Chyou PH, Kagan A, Nomura AMY, Yano K. Serum cholesterol and mortality among Japanese-American men. The Honolulu (Hawaii) heart program. *Arch Intern Med* 1991;151:969-72.
- 25 Iso H, Jacobs DR Jr, Wentworth D, Neaton JD, Cohen JD. Serum cholesterol levels and six-year mortality from stroke in 350977 men screened for the multiple risk factor intervention trial. *N Engl J Med* 1989;320:904-10.
- 26 Dyer AR, Stamler J, Paul O, Shekelle RB, Schoenberg JA, Berkson DM, *et al*. Serum cholesterol and risk of death from cancer and other causes in three Chicago epidemiological studies. *J Chronic Dis* 1981;34:249-60.
- 27 Chen Z, Collins R, Peto R. Serum cholesterol levels and stroke mortality. *N Engl J Med* 1989;321:1339-41.
- 28 Reed DM. The paradox of high risk of stroke in populations with low risk of coronary heart disease. *Am J Epidemiol* 1990;131:579-88.
- 29 Wolf PA, Kannel WB, Dawber TR. Prospective investigation: the Framingham study and the epidemiology of stroke. *Adv Neurol* 1978;19:107-20.
- 30 Westlund K, Nicolaysen R. Ten-year mortality and morbidity related to serum cholesterol. *Scand J Clin Lab Invest* 1972;suppl 127:1-24.
- 31 Aronow WS, Gutstein H, Lee NH, Edwards M. Three-year follow-up of risk factors correlated with new atherothrombotic brain infarction in 708 elderly patients. *Angiology* 1988;39:563-6.
- 32 Shimamoto T, Komachi Y, Inada H, Doi M, Iso H, Sato S, *et al*. Trends for coronary heart disease and stroke and their risk factors in Japan. *Circulation* 1989;79:503-15.
- 33 Ueshima H, Iida M, Shimamoto T, Konishi M, Tsujioka K, Tanigaki M, *et al*. Multivariate analysis of risk factors for stroke. Eight year follow-up study of farming village in Akita, Japan. *Prev Med* 1980;9:722-40.
- 34 Wilhelmssen L. Synergistic effect of risk factors. *Clin Exp Hypertens [A]* 1990;12:845-63.
- 35 Welin L, Svärdsudd K, Wilhelmssen L, Larsson B, Tibblin G. Analysis of risk factors for stroke in a cohort of men born in 1913. *N Engl J Med* 1987;317:521-6.
- 36 Davis CE, Rifkind BM, Brenner H, Cordon DJ. A single cholesterol measurement underestimates the risk of coronary heart disease. An empirical example from the lipid research clinics mortality follow-up study. *JAMA* 1991;264:3044-6.
- 37 Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika* 1982;69:331-42.
- 38 Kagan A, Harris BR, Winkelstein W Jr, Johnson KG, Kato H, Syme SL, *et al*. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: demographic, physical, dietary and biochemical characteristics. *J Chronic Dis* 1974;27:345-64.
- 39 Yano K, Reed DM, MacLean CJ. Serum cholesterol and hemorrhagic stroke in the Honolulu heart program. *Stroke* 1989;20:1460-5.
- 40 Mui K. Relationship of risk factors to subsequent development of stroke and ischemic heart disease in a rural community. *Osaka City Med J* 1989;35:145-61.
- 41 Gatchev O, Råstam L, Lindberg G, Gullberg B, Eklund GA, Isacson S-O. Subarachnoid hemorrhage, cerebral hemorrhage and serum cholesterol concentration in men and women. *Ann Epidemiol* 1993;3:403-9.
- 42 Bamford J, Sandercock P, Dennis M, Warlow C. A prospective study of acute cerebrovascular disease in the community: the Oxfordshire community stroke project—1981-86. *J Neurol Neurosurg Psychiatry* 1990;53:16-22.
- 43 Strachan D, Rose G. Strategies of prevention revisited: effects of imprecise measurement of risk factors on the evaluation of "high-risk" and "population-based" approaches to prevention of cardiovascular disease. *J Clin Epidemiol* 1991;44:1187-96.
- 44 Taggart H, Stout RW. Reduced high density lipoprotein in stroke: relationship with elevated triglyceride and hypertension. *Eur J Clin Invest* 1979;9:219-21.
- 45 Kannel WB. Epidemiology of cerebrovascular disease. In: Russel WWT, ed. *Cerebral disease*. Edinburgh: Churchill Livingstone, 1976:1-23.
- 46 Ryu JE, Howard G, Craven TE, Bond MG, Hagamann AP, Crouse Jr. Postprandial triglyceridemia and carotid atherosclerosis in middle-aged subjects. *Stroke* 1992;23:823-8.
- 47 Bihari-Varga M, Székely J, Gruber E. Plasma high density lipoproteins in coronary, cerebral and peripheral vascular disease: the influence of various risk factors. *Atherosclerosis* 1981;40:337-45.
- 48 Murai A, Tanaka T, Miyahara T, Kameyama M. Lipoprotein abnormalities in the pathogenesis of cerebral infarction and transient ischemic attack. *Stroke* 1981;12:167-72.
- 49 Gordon T, Kannel WB, Castelli WP, Dawber TR. Lipoproteins, cardiovascular disease and death: the Framingham study. *Arch Intern Med* 1981;141:1128-31.
- 50 Stampfer MJ, Sacks FM, Salvini S, Willett WC, Hennekens CH. A prospective study of cholesterol, apolipoproteins and the risk of myocardial infarction. *N Engl J Med* 1991;325:373-81.
- 51 Yatsu FM, Fisher M. Atherosclerosis: current concepts on pathogenesis and interventional therapies. *Ann Neurol* 1989;26:3-12.

(Accepted 14 April 1994)

Multicentre criterion based audit of the management of induced abortion in Scotland

Gillian C Penney, Anna Glasier, Allan Templeton

Abstract

Objectives—To assess and improve the quality of care provided to women undergoing induced abortion.

Design—Two rounds of prospective, criterion based case note review audit.

Setting—Ten NHS gynaecology units throughout Scotland.

Subjects—2004 patient episodes of abortion care identified consecutively during two rounds of audit. The first round comprised 967 cases and the second round 1037.

Interventions—Dissemination of results from the first round of audit and recommendations for change in the form of a written report and at postgraduate meetings in participating hospitals.

Main outcome measures—Improvements in quality of care as assessed against 16 previously agreed criteria, both overall across the 10 study hospitals and within individual hospitals.

Results—Overall, four significant improvements occurred: increased availability of early medical abortion, decreased utilisation of surgical abortion at very early gestation, increased use of mifepristone priming before second trimester medical abortion, and increased provision of follow up. At the individual hospital level 42 of 150 elements of care studied were "close to optimal" at the time of the first round of audit, rising to 54 at the second round (NS). A total of 31 significant improvements in individual elements of care occurred, but 11 significant deteriorations also occurred (at the $P < 0.05$ level).

Conclusions—The prospective multicentre audit proved feasible and achieved the aims of any form of audit in terms of identifying deficiencies and variations in care. The audit results prompted objective review of local abortion services in participating hospitals. At least for some elements of care in some hospitals significant improvements were detectable.

Introduction

Induced abortion is one of the commonest components of the Scottish NHS gynaecological workload, accounting for around 11 000 procedures per year.¹ Abortion care was therefore considered to be a particularly appropriate topic for medical audit and was recently addressed by means of a criterion based approach in the gynaecology audit project in Scotland.²

Methods

A list of criteria for good quality care was agreed by a combination of objective review of contemporary medical publications, panel discussions, and postal survey of all consultant gynaecologists in Scotland (response rate 92%), as described.³ The 16 criteria which were addressed in the multicentre audit are listed in the box.

Ten hospitals throughout Scotland representing 10 different health board areas and employing around half of all consultant gynaecologists in Scotland were included. During each of two audit periods an audit

Department of Obstetrics and Gynaecology, Aberdeen Maternity Hospital, Aberdeen AB9 2ZD
Gillian C Penney, *audit research fellow*
Allan Templeton, *professor*

Lothian Health Board, Family Planning and Well Woman Services, Edinburgh EH4 1NL
Anna Glasier, *director of family planning and well woman services*

Correspondence to: Dr Penney.

BMJ 1994;309:15-9

Criteria for good quality care addressed in case note review audit and agreed among consultant gynaecologists in Scotland

Organisation of service

- (1) All women requesting abortion should be offered appointment with gynaecologist within five days of referral
- (2) Induced abortion should be undertaken within seven days of appointment with gynaecologist
- (3) In absence of specific medical or social contraindications women undergoing abortion should be managed as day cases

Preabortion investigations

- (4) Woman's Rh status should be ascertained and Rh prophylaxis given after abortion if indicated
- (5) Gestation need not routinely be confirmed by ultrasound scan
- (6) Cervical smear history should be recorded for all women undergoing abortion
- (7) Women undergoing abortion should be screened for genital tract infection and treated if indicated

Methods of abortion

- (8) Women presenting early in pregnancy (<9 weeks' gestation) should have option of surgical or medical abortion
- (9) Suction termination should not be undertaken in very early pregnancy (≤ 6 weeks' gestation) because of high incidence of failure
- (10) In selected women undergoing surgical abortion, technique for cervical predilatation should be used
- (11) All surgical abortions at gestations >10 weeks should be carried out by experienced gynaecologist (post-MRCOG or equivalent)
- (12) For second trimester medical abortions pretreatment with mifepristone reduces induction abortion interval

Postabortion care

- (13) Before being discharged each patient should have agreed a contraceptive plan
- (14) Before being discharged each patient should be offered contraceptive supplies
- (15) A follow up appointment, either at hospital or with referring doctor, should be given to every patient
- (16) Follow up appointment should be within two weeks of abortion procedure

assistant (with a medical secretarial background) in each hospital identified all patients undergoing induced abortion. Mothers undergoing abortion because of fetal abnormality were excluded as they were thought to raise different care issues. Data relating to the agreed criteria were collected in a standardised manner in all 10 hospitals by transfer of information from case notes on to a review document. Each audit assistant was trained by the research fellow (GCP) in data transfer. Data from all hospitals were entered into a purpose designed database by using Paradox (Borland) software on an IBM compatible personal computer. Statistical significance testing by χ^2 analysis (unless otherwise indicated) was performed with Instat 2 (Graphpad) software.

After the first audit period results and recommenda-

tions for change were disseminated as a written report mailed to all consultant gynaecologists. In addition, in seven of the 10 hospitals a postgraduate meeting was held, allowing a further opportunity to highlight local issues.

Results

CASES IDENTIFIED

A data collection period of six months (beginning in October 1992) was available for the first round of audit. It was planned to review 100 consecutive cases managed in each hospital during this time. In the larger hospitals 100 cases were identified during as little as six weeks, whereas in one of the smaller hospitals only 67 cases presented during the whole six months. Thus in total 967 cases were studied. Owing to constraints of the project timetable and funding, data collection for the second round of audit was restricted to a two month period (beginning in August 1993). All cases managed in each hospital during this time were identified and reviewed. The mean number of cases per hospital was 104 (range 20 to 266), and in total 1037 cases were studied.

OVERALL IMPROVEMENTS

At the time of the second round of audit four significant overall improvements across the 10 hospitals were detected:

- The use of medical abortion for women at <9 weeks' gestation (criterion 8) rose from 39 to 516 (7.6%) to 172 of 541 (31.8%) ($P < 0.0001$)
- The inappropriate use of surgical abortion in women at <7 weeks' gestation (criterion 9) decreased from 68 to 85 (80.0%) to 56 of 98 (57.1%) ($P = 0.0017$)
- The use of mifepristone cervical priming before midtrimester medical abortion (criterion 12) increased from 15 of 64 women (23.4%) to 64 of 102 (62.7%) ($P < 0.0001$)
- The recording of a follow up arrangement in the case notes (criterion 15) increased from 52% to 69% of cases (Wilcoxon signed rank test, $P = 0.037$), and the advising of follow up within the recommended interval of two weeks after abortion also increased (from 5% to 32%; Wilcoxon signed rank test, $P = 0.0645$).

There were no overall deteriorations in relation to any elements of care.

IMPROVEMENTS IN INDIVIDUAL HOSPITALS

Fifteen of the agreed criteria were addressed on an individual basis in each of the 10 hospitals. (Criterion 9 was omitted in view of the small numbers of patients presenting at ≤ 6 weeks in any one hospital.) Thus 150 individual elements of care were assessed. Results from

TABLE 1—First round of audit. Results of individual hospitals in relation to 15 of agreed audit criteria

Criterion No	Element of care	Hospital No (No of cases studied)									
		1 (n=100)	2 (n=100)	3 (n=100)	4 (n=100)	5 (n=100)	6 (n=100)	7 (n=100)	8 (n=100)	9 (n=100)	10 (n=67)
1	% Of cases seen within five days	30	29	86	52	42	64	40	55	30	54
2	% Of procedures done within seven days	78	89	89	86	93*	70	63	67	92*	89
3	% Of patients managed as day cases	89	63	93*	95*	84	91*	70	7	72	48
4	% Of cases where Rh status was recorded	95*	99*	98*	100*	98*	99*	100*	86	94*	97*
5	% Of cases having ultrasound scan	33	46	17*	17*	25*	37	97	83	13*	9*
6	% Of cases where cervical smear history was recorded	63	21	55	0	77	77	84	64	0	12
7	% Of cases where genital tract swabs were taken	92*	14	2	1	2	88	15	1	0	0
8	Early medical abortion service available (yes (Y), no (N))	Y*	Y*	N	N	Y*	N	N	N	N	N
10	% Of high risk surgical abortions having predilatation	61	69	25	71	92*	10	13	62	10	87
11	% Of surgical abortions at >10 weeks performed by senior registrar or above	67	54	57	87	100*	100*	100*	79	100*	100*
12	Mifepristone used in second trimester medical abortion (yes (Y), no (N))	Y*	Y*	N	N	N	N	N	N	N	N
13	% Of cases where contraceptive plan was recorded	95*	57	100*	96*	98*	98*	88	98*	96*	94*
14	% Of cases where contraceptive supplies were given by hospital	29	16	43	24	83	20	16	12	80	81
15	% Of cases where follow up arrangements were recorded	74	25	100*	99*	91*	7	37	55	26	6
16	% Of cases where follow up within two weeks was advised	13	4	18	0	1	3	0	4	6	0

*Asterisk indicates element of care where performance was "close to optimal."

TABLE II—Second round of audit. Results of individual hospitals in relation to 15 of agreed audit criteria

Criterion No	Element of care	Hospital No (No of cases studied)									
		1 (n=266)	2 (n=209)	3 (n=149)	4 (n=112)	5 (n=49)	6 (n=77)	7 (n=81)	8 (n=41)	9 (n=33)	10 (n=20)
1	% Of cases seen within five days	66↑	26	89	17↓	33	29↓	53	49	67↑	5↓
2	% Of procedures done within seven days	78	86	84	77	75↓	75	75	73	85	80
3	% Of patients managed as day cases	85	82↑	96*	86↓	94*	90*	72	27↑	82	65
4	% Of cases where Rh status was recorded	98*	95*	95*	100*	100*	99*	100*	98*	91*	95*
5	% Of cases having ultrasound scan	29	67↓	13*	22*	31	25*	100	98↓	18*	15*
6	% Of cases where cervical smear history was recorded	81↑	10↓	38↓	99*↑	96*↑	70	76	80	51↑	55↑
7	% Of cases where genital tract swabs were taken	92*	6	0	12	86↑	88	10	0	0	0
8	Early medical abortion service available (yes (Y), no (N))	Y*	Y*	Y*↑	N	Y*	N	Y*↑	N	N	N
10	% Of high risk surgical abortions having predilatation	77	86	53	74	100*	44↑	0	22	67	100*
11	% Of surgical abortions at > 10 weeks performed by senior registrar or above	59	75	33	100*	100*	100*	100*	57	100*	100*
12	Mifepristone used in second trimester medical abortion (yes (Y), no (N))	Y*	Y*	Y*↑	N	Y*↑	Y*↑	Y*↑	N	N	N
13	% Of cases where contraceptive plan was recorded	92*	46	99*	100*	100*	99*	95*	100*	97*	85
14	% Of cases where contraceptive supplies were given by hospital	42↑	6↓	52	44↑	96*↑	18	26	17	97*↑	50↓
15	% Of cases where follow up arrangements were recorded	89↑	52↑	94*	100*	98*	14	80↑	98*↑	12	55↑
16	% Of cases where follow up within two weeks was advised	59↑	35↑	19	93*↑	2	0	2	66↑	3	40↑

*Asterisk indicates element of care where performance was "close to optimal."

↑Indicates significant improvement ($P < 0.05$).↓Indicates significant deterioration ($P < 0.05$).

the first round of audit are summarised in table I. Of these 150 elements of care, 42 (28.0%) were judged to be "close to optimal" (arbitrarily defined as $\geq 90\%$ of cases meeting the criterion, or (for criterion 5) $\leq 25\%$ of cases undergoing ultrasound scanning, (for criteria 8 and 12) mifepristone being available).

Performance was particularly good in relation to recording of Rh status (nine hospitals achieved $> 90\%$) and recording of patient's contraceptive plan (eight hospitals achieved $> 90\%$). No hospital achieved close to optimal performance in relation to seeing patients within five days of referral, recording of cervical smear history, providing contraceptive supplies, or suggesting follow up within two weeks of the abortion.

The results from the second round of audit are summarised in table II. Fifty four elements of care (36.0%) were judged to be close to optimal. This increase over the first round failed to reach significance ($P=0.1732$). However, major improvements occurred in relation to 31 individual elements of care. These comprised the introduction of a new service (either early medical abortion or the use of mifepristone for second trimester medical abortion) in six instances or significant ($P < 0.05$) improvements in the proportion of cases meeting a given criterion in 25 instances. The most common significant improvements were in relation to recording of cervical smear history, recording of a follow up plan, and advising early follow up (each occurred in five hospitals).

Eleven significant deteriorations in the proportion of cases meeting a given criterion also occurred. The most common deterioration was in relation to the proportion of cases seen by a gynaecologist within five days of referral (three hospitals).

There was no significant difference between the number of improvements occurring in the seven hospitals in which a postgraduate meeting was held (23 improvements out of a possible 105) and the three hospitals in which results and recommendations were disseminated as a report only (eight out of a possible 45) ($P=0.6632$).

Discussion

The multicentre audit proved feasible and achieved good cooperation among gynaecologists. While acknowledging the theoretical advantages of clinical rather than medical audit we perceived that this project engendered a strong sense of ownership as a result of being a confidential single discipline exercise initiated by the Scottish branch of the Royal College of Obstetricians and Gynaecologists, of which all participants were members.

The first round of audit enabled variations and deficiencies in care to be identified and undoubtedly

prompted open and objective discussions about local abortion services in participating hospitals. It might be argued that some of the changes detected in the second round were unrelated to the audit feedback exercise. However, one of the most dramatic improvements (the increase in mifepristone priming before midtrimester abortion) occurred at a time unrelated to new publications on the topic and before the granting of a product licence. Medical audit can never be the sole stimulus to change; it acts alongside other educational initiatives. In this instance, however, we believe the audit was instrumental in changing practice.

It is clear that in order to maximise the benefits from the national investment in clinical audit attention must be directed at identifying the most effective strategies for disseminating audit results and for implementing change. Russell and Grimshaw have provided suggestions for the effective dissemination and implementation of clinical practice guidelines⁴ (to which audit recommendations are closely related).⁵ They suggest that dissemination is best accomplished by "specific educational intervention" rather than simply "mailing targeted groups." However, in this study no more improvements occurred in those hospitals where a postgraduate meeting was held specifically to address the audit findings than in those where a postal report only was provided.

Implementation entails encouraging clinicians to adopt the audit recommendations. The only implementation strategy employed in this study was fore-

Clinical implications

- The quality of care provided to women requesting abortion was assessed by means of criterion based audit in hospitals throughout Scotland
- Deficiencies in care and variations among hospitals were detected
- Participating clinicians responded to the audit results and to recommendations made by the audit team
- Medical audit can act along with other educational initiatives in increasing awareness of effective treatments and in providing a stimulus to change
- This audit exercise prompted open and objective discussions among clinicians about many aspects of patient care and is believed to have been instrumental in stimulating significant changes in relation to some

warning participants that a second round of audit was imminent. A more effective strategy might have been to incorporate recommendations in a structured patient record document for use while managing abortion referrals, a technique which has proved successful in implementing guidelines for the management of infertility.⁶

The modest success of this national audit has confirmed an interest and willingness to participate in medical audit among Scottish gynaecologists. Refinements in audit methodology are required to translate this good will into substantial improvements in patient care.

We acknowledge the cooperation of the following consultant gynaecologists/audit assistants who coordinated the audit in their respective hospitals: Aberdeen Royal Hospitals NHS Trust, P B Terry/J Mowat, J Fryers; Edinburgh Royal Infirmary, G E Smart/K Duncan; Ninewells Hospital, Dundee, N Patel/W Barrie, J Watson; Southern General Hospital, Glasgow, M Carty/C Doak; Ayrshire Central Hospital, S Prigg/J Morland, A-M Garvey, J Howie, C McIntyre; Raigmore Hospital, Inverness, M Hulse/M

Gordon; Victoria Hospital, Kirkcaldy, M Hill/M Telfer; Dumfries and Galloway Infirmary, G Gordon/A Graham; Inverclyde Royal Hospital, Greenock, L Cassidy/U Stevenson; Hairmyres Hospital, East Kilbride, J Grant/M McCartney. Elaine Stirton computerised the data and typed the manuscript. Daphne Russell provided statistical advice. The gynaecology audit project in Scotland is funded by a grant from the Clinical Resource and Audit Group of the Scottish Office of the Home and Health Department (but the views expressed are ours and not necessarily those of the Clinical Resource and Audit Group or the department).

1 Abortion statistics 1991. *Health Briefing* 1992;92:04.

2 Shaw CD. Criterion based audit. In: Smith R, ed. *Audit in action*. London: BMJ, 1992;122-8.

3 Penney GC, Glasier A, Templeton A. Agreeing criteria for audit of the management of induced abortion: an approach by national consensus survey. *Quality in Health Care* 1993;2:167-9.

4 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317-22.

5 Clinical Resource and Audit Group (SOHHD). *Clinical guidelines: report of a working group*. Edinburgh: Clinical Resource and Audit Group, 1993.

6 Emslie C, Grimshaw J, Templeton A. Do clinical guidelines improve general practice management and referral of infertile couples? *BMJ* 1993;306:1728-31.

(Accepted 18 April 1994)

Commentary

Validity of assessing change through audit

Ken MacRae

The difference between audit and research was neatly encapsulated by the subheading to a leading article by Smith: "Research is concerned with discovering the right thing to do; audit with ensuring that it is done right."¹

Penney *et al* have reported an audit of the management of induced abortion in 10 hospitals in Scotland, comparing the results obtained in a first round of six months in 1992 with those obtained in a second round of two months in 1993.² The authors conclude that four statistically significant improvements were detectable when all 10 hospitals were taken together and that no overall deteriorations occurred in any elements of care. A rather different picture emerged when individual hospitals were considered. Considering 15 of the 16 criteria of care, giving 150 comparisons of the two rounds of audit, it was concluded that major improvements occurred in 31 comparisons and significant deteriorations occurred in 11. Despite acknowledging that some of the improvements recorded might have been unrelated to the audit feedback exercise, Penney *et al* believe that audit was instrumental in producing the dramatic improvement in the frequency of mifepristone priming before midtrimester abortion.

The problem with drawing any such conclusion from such evidence is not whether it is true but whether it is valid. Methodologically, a comparison of two audit rounds is like a clinical trial using historical controls. In 1977 Pocock addressed the temptation to draw conclusions from non-randomised comparisons in clinical trials when no sources of bias can be identified.³ He reviewed a number of cancer chemotherapy trials in the United States and found 19 pairs carried out by cooperative groups in which the same control treatment had been used in consecutive randomised trials. Most of the trials had over 100 patients per treatment, and no obvious sources of bias suggesting differences between the trials in each pair were identifiable. Despite this, the death rates with the

same treatment in the two successive trials showed changes ranging from a decrease of 46% to an increase of 24%. Four of the 19 pairs showed changes significant at the 5% level. Pocock concluded: "Such marked evidence of differences between trials indicates that any comparison of treatments not within an RCT [randomised controlled trial] must be deemed highly suspect."³

Logic of a clinical trial

It is worth reviewing the logic on which drawing a causal inference is based in a clinical trial. An apparent difference between two treatments has four possible explanations, but these are not mutually exclusive. Firstly, the difference could be "real," indicating a difference due to the difference in treatment. Secondly, it may simply have occurred by chance even though the two treatments are equally effective (or ineffective). Thirdly, the patients in the two treatment groups might be different—that is, there might be an allocation or group membership bias. Fourthly, there could be a difference in the way that the outcome of treatment was judged—that is, an assessment or measurement bias.

The logic of a clinical trial is that we can conclude that a difference is real only if we can eliminate the other three explanations. The role of chance is considered by performing a statistical test to give a P value. The usual convention is that if P is less than 0.05 chance may be rejected as a reasonable explanation of the difference. A particular problem arises when several P values have been calculated, as it becomes ever more likely that low values will occur even if the true differences in all the comparisons made are zero. Though a low P value is necessary in order to draw a cause and effect conclusion, it is not enough. Allocation and assessment bias also need to be considered. Allocation bias is addressed usually by allocating patients randomly to the treatment groups. Even if the

Charing Cross and Westminster Medical School, London W6 8RP
Ken MacRae, reader in medical statistics

Correspondence to:
5 Northcroft Terrace,
Northcroft Road, London
W13 9SP