An Analysis of Randomized Controlled Trials Published in the US Family Medicine Literature, 1987–1991

Christopher A. Silagy, MBBS, PhD, FRACGP, FAFPHM; David Jewell, MA, MB, BChir, MRCGP; and David Mant, MBChB, MSc, MRCGP, MFPHM

Adelaide, Australia, and Bristol and Southampton, United Kingdom

Background. Randomized controlled trials (RCTs) are frequently used to evaluate the effectiveness of health care interventions in family medicine. The value of the information obtained from RCTs depends largely on the quality of design and the way in which they are conducted and reported. Despite the increasing number of RCTs being conducted in family medicine, there is a scarcity of descriptive data on the methodological characteristics, including design features and quality of RCTs in this setting.

Methods. All 55 RCTs published in four peer-reviewed US family medicine journals between 1987 and 1991 were identified and their methodological characteristics reviewed. Three potential sources of bias were assessed in each of the trial reports: (1) control of selection bias at entry, (2) control of selection bias after entry, and (3) control of bias in assessing outcome(s).

Results. Fifty-five RCTs published between January 1987 and December 1991 were identified in the four journals. The absolute number of RCTs published over the 5 years increased steadily, and there was a 49% increase in the proportion of RCT articles. Measures used to control for selection bias before entry into the study were reported in 14 (25%) of the RCTs, the statistical power of the trial in 5 (9%), and whether the study had been reviewed by an institutional review board in 6 (11%).

Conclusions. The RCTs analyzed offered some imaginative solutions to the logistic difficulties of conducting RCTs in general practice. Nevertheless, the methodology and reporting of RCTs in the future should be improved.

Key words. Clinical trials; family practice; research design; clinical protocols; randomized controlled trials; peer review, research.

(*J Fam Pract 1994; 39:236-242*)

Randomized controlled trials (RCTs) are commonly regarded as the "gold standard" research design for evaluating the effectiveness of health care interventions.[1] An increasing number of trials conducted in family medicine settings are being published and are directly relevant to family medicine.[2] Historically, RCTs focused on the evaluation of new pharmacological agents. More recently, the same research method is being used to evaluate nonpharmacological treatments, such as physical therapies, counseling, educational strategies, and interventions concerned with the structure and function of primary care as a health service, eg, cost-reduction strategies and the use of computer reminders to increase the level of preventive activity.

Furthermore, there is a trend toward promoting more "pragmatic" clinical trials that focus on the effectiveness of an intervention in the "real world" as opposed to "explanatory" trials that assess the efficacy of a particular intervention under much more stringent research conditions.[1]

The value of information obtained from an RCT depends largely on the quality of the design and the way in

Submitted, revised, May 31, 1994.

From the Department of General Practice, Flinders University of South Australia School of Medicine, Adelaide, Australia (C.A.S.); the Department of Epidemiology and Public Health Medicine, University of Bristol, Bristol (D.J.), and the Department of Primary Health Care, University of Southampton, Southampton (D.M.), United Kingdom. Requests for reprints should be addressed to Professor C.A. Silagy, Department of General Practice, Flinders University of South Australia School of Medicine, PO Box 2100, Adelaide SA 5001, Australia.

ISSN 0094-3509

which it is conducted and reported. Several papers have examined logistic aspects associated with individual RCTs.[3–7] In a previous historical analysis of RCTs that covered a 39-year period,[3] we made no attempt to comment on the methodological aspects of the trials because of the wide span of time involved.

The aim of this study was to describe the methodological features of RCTs that have been reported in the four main research-based family medicine journals published in the USA—*The Journal of Family Practice, The Journal of the American Board of Family Practice, Family Medicine,* and *Family Practice Research Journal*—and to assess the quality of the trials. The assessment criteria were in keeping with accepted standards for reporting clinical trials.[8] In assessing trial quality, we focused on control of bias in three areas: (1) selection bias at entry, (2) selection bias after entry, and (3) bias in assessment of outcome(s). Since the standards for conducting and reporting RCTs have advanced considerably during the past decade,[1] we confined the study to a recent 5-year period (1987 to 1991, inclusive) during which the publication standards remained fairly constant.

Methods

Inclusion Criteria

Each trial had to meet the following two methodological criteria to be eligible for inclusion in the study: (1) it must have included at least two groups, and (2) allocation to the groups must have been either by formal randomization or by a quasi-random method (eg, alternation).

Identification of Trials

Fifty-five RCTs were identified by means of a manual search of all issues of the four journals published between January 1987 and December 1991, inclusive. Each issue was individually scrutinized by an experienced family medicine researcher to identify RCTs that met the inclusion criteria. A full list of the 55 trials included in this analysis is available from the authors on request.

Extraction of Information

The authors of the current study each reviewed one third of the identified trials and extracted and coded the information listed in Table 1. An assessment was made of interobserver reliability for assessment of study quality. The authors discussed any uncertainties about the data extraction or coding for studies. The subject area addressed in each trial was coded using the International Classification of Health Care Problems in Primary Care (ICHCPPC),[9] to which a few supplementary categories were added to cover RCTs involving medical education and health services research, for which there were no codes.

Table 1. Information Required for the Assessment of a Randomized Controlled Trial

1. Country of origin
2. Primary site of the trial (single general practice, multiple practices, community setting, hospital)
3. Subject area
4. Study population (volunteers, screened subjects, all eligible patients)
5. Trial inclusion and exclusion criteria
6. Study design (method of allocation, unit of allocation, blinding)
7. Structure (parallel group, crossover, factorial design)
8. Number, nature, and duration of intervention(s) and control groups
9. Analysis ("intention to treat" vs "on randomized treatment")
10. Withdrawals
11. Adverse effects
12. Ethics approval
13. Use of invasive investigations (eg, blood tests)
14. Use of practice nurses or research assistants, nurses, or both.

Quality Assessment

The methodological quality of each trial was assessed using a simplified scheme (Table 2) previously developed by Chalmers et al.[10] It involves assessing three dimensions of trial methodology that are important potential sources of significant bias: (1) the quality of the random allocation (ie, control of selection bias at entry), (2) the extent to which the primary analysis included every person entered into the randomized cohorts (ie, control of selection bias after entry), and (3) the extent to which those assessing outcomes were kept unaware of the group assignment of the individuals examined (ie, control of bias in assessing outcomes). For each of the three dimensions, we used a three-point rating scale, ranging from a score of 3, indicating maximal effort to control potential bias, to 1, indicating little or no effort to control potential bias.

Data Analysis

Essential frequencies were calculated for the number of RCTs in the different subject areas and for the number that met the various methodological characteristics examined. The frequency with which various subjects areas were addressed in the clinical trials included in this analysis was compared with that of primary care RCTs from the same period. The primary care RCTs were identified by means of a MEDLINE search and are the basis of a developing international register of RCTs in primary

Table 2. Criteria for the Assignment of Scores in the Assessment of Bias Control in a Randomized Controlled Trial

A. Control of selection bias at entry
"Could those enrolling the study participants know which treatment was next in line?"
3 points: Random treatment by telephone communication or indistinguishable drug treatments precoded by a pharmacy
2 points: Use of less secure methods of reducing advance knowledge of treatment allocation (eg, sealed envelopes)
1 point: No attempt at blinding treatment allocation (eg, use of patient record numbers, dates of birth, day of the week), or no details of the random allocation procedure provided in the report

B. Control of postselection bias
"Was the primary analysis based on all cases allocated ro receive one or other of the alternative treatments?"
3 points: Primary analysis based on all participants who were randomized, irrespective of whether they withdrew from the trial (ie, an "intention to treat" approach to analysis)
2 points: Exclusions after randomization too few to introduce bias in assessing the principal outcome, and/or loss to follow-up independent of the allocated treatment
1 point: Trials that cannot be assigned either two or three points

C. Control of outcome assessment bias
"To what extent was assessment of the principal outcome likely to have been affected by knowledge of the treatment allocation?"
3 points: Adequate precautions taken to blind those assessing the principal outcome, or principal outcome unambiguous and blinding irrelevant
2 points: Some precautions taken to blind those assessing the principal outcome, but inadequate to exclude the possibility of bias
1 point: No precautions taken to minimize bias in assessing the principal outcome

care.[2] For nonparametric comparisons, a chi-square test was used, and a two-tailed *P* value <.05 was regarded as statistically significant. Ninety-five percent confidence intervals (CI) of proportions were calculated using the CIA software program.[11]

Results

A total of 55 RCTs were identified in the issues of the four journals published between January 1987 and December 1991: 31 in *The Journal of Family Practice,* 10 in *Family Medicine,* 9 in *The Journal of the American Board of Family Practice,* and 5 in *Family Practice Research Journal.* This represents a steady increase in the absolute number of RCTs published over the past 5 years and a 49% increase in the proportion of RCTs to the total number of primary care articles (Figure). It is difficult to make inter-journal comparisons because of the small number of RCTs published in each of the four journals and the varying frequency of publication (from monthly to quarterly).

All the RCTs were published as full reports except for one, which reported only the baseline characteristics of participants in a trial following completion of the recruitment phase[12]; and all were carried out in the United States except for two, which were from Denmark and Canada.

Thirty-six (65%) of the trials were carried out in a primary care setting; 4 (7%) in other community settings, 3 (5%) were based in hospitals, and 7 (13%) were in a combination of these settings. Five of the articles did not state the primary site in which the trial was carried out.

Nature of Interventions

Thirty-three of the trials (60%) involved comparison of a single intervention and a control group; 16 (29%) involved comparison of two active interventions, and 6 (11%) involved comparison of three or more active interventions. Twenty-three (42%) of the interventions were

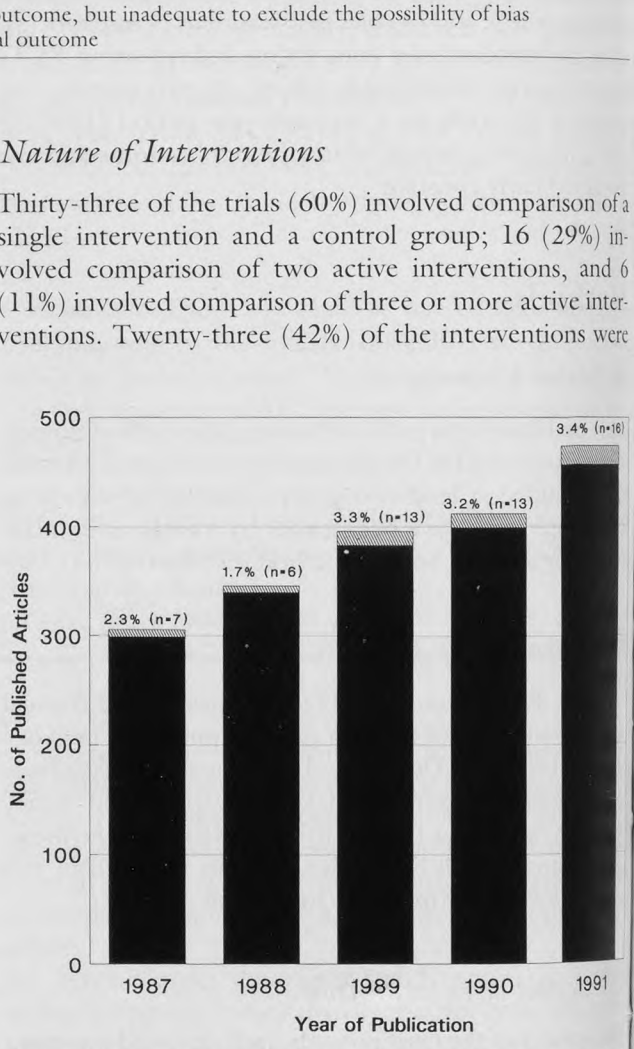

Figure. The number of randomized controlled trials (hatched bars) and other articles (solid bars) published annually between 1987 and 1991 in US family medicine journals. The number of randomized controlled trials published annually and the percentage of the total primary care articles this number represents are shown at the top of each bar.

Table 3. Subject Areas Covered by Randomized Controlled Trials (RCTs) in Family Medicine

Subject Area	RCTs Published in Four US Primary Care Journals, % (95% CI) (n=55)	RCTs in International Register of RCTs,* % (n=266)
ICHCPPC Classification†		
Infective and parasitic diseases	0 (0.0–6.4)	2.3
Neoplasms	1.8 (0.0–9.6)	0.4
Endocrine, nutritional, metabolic diseases	7.3 (2.0–17.3)	3.7
Blood diseases	0 (0.0–6.4)	0.4
Mental diseases	10.9 (4.0–21.9)	17.7
Nervous system, sense organ diseases	1.8 (0.0–9.6)	4.2
Circulatory system diseases	9.1 (3.0–19.6)	18.8
Respiratory system diseases	12.7 (5.2–24.1)	14.7
Digestive system diseases	1.8 (0.0–9.6)	3.8
Genitourinary system diseases	9.1 (3.0–19.6)	8.6
Pregnancy, childbirth, puerperium	1.8 (0.0–9.6)	2.6
Skin, subcutaneous tissue diseases	3.6 (0.4–12.1)	1.5
Musculoskeletal diseases	0 (0.0–6.4)	6.0
Perinatal morbidity and mortality	0 (0.0–6.4)	0.8
Supplementary Classification†		
Preventive medicine	25.5 (14.7–39.0)	6.4
Administrative procedures	0 (0.0–6.4)	0.8
Social problems	0 (0.0–6.4)	0.4
Miscellaneous	1.8 (0.0–9.6)	0.4
Other‡		
Health service research	1.8 (0.0–9.6)	4.1
Medical education	10.9 (4.0–21.9)	2.6

**Identified via MEDLINE search of RCTs published 1987–1991.[3]*
†Adapted from the International Classification of Health Care Problems in Primary Care (ICHCPPC).[9]
‡Not coded in the ICHCPPC.

nonpharmacological, 20 (36%) involved a pharmacological agent, and 12 (22%) compared the provision of one or more different health services.

Placebo controls were used in 11 (55%) of the trials involving a pharmacological therapy. In the remainder, the efficacy of a new pharmacological regimen was compared with that of an existing one. Among trials of nonpharmacological therapies or health-service interventions, the control group reflected "usual care" in 22 cases (63%).

The main subject areas addressed by the RCTs published in the four journals are shown in Table 3. During the same period, there was a higher proportion of trials involving preventive and medical education interventions in these four journals than among those included in the international register of RCTs, which is compiled from a comprehensive MEDLINE search. Forty-three (78%) of the trials were of relatively short duration, involving less than 6 months of intervention and follow-up. Only one study had a treatment and follow-up period that extended beyond 12 months.

Design Aspects

In 38 (69%) of the trials, participants were recruited on a nonselective basis. In 12 others (22%), the study population was derived by a screening process established specifically for the purpose of identifying eligible participants. Only 4 (7%) of the trials specifically sought volunteers, and one made no mention of how the study population was obtained. Description of inclusion and exclusion criteria were frequently incomplete.

Patients, rather than physicians or their practices, were allocated to receive intervention in 48 (87%) of the trials. In one trial, which involved topical applications for wound healing, patients were allocated to receive the different treatments simultaneously but applied randomly to different parts of the body. Of the remaining seven trials, the unit of randomization was physicians in six and practices in one.

Forty-five (82%) of the reports stated that allocation to the intervention and control groups was by randomization. In eight studies, a systematic procedure, such as alternation, date of birth, or days of the week, was used to determine allocation to the different trial groups. Of the remaining two studies, one used a combination of systematic allocation to different treatment groups and nonrandom allocation to a control group. The other study did not formally state the method of allocation.

Blinding was used in just over one half of the trials: 13 (24%) were double-blind, and 17 (31%) were single-blind. The remaining 25 trials were conducted without blinding. Forty-nine (89%) of the trials were conducted using a parallel group design. A crossover design was used in only six (11%) of the studies.

Power calculations were reported in five trials. Although these calculations were not checked and no calculations were made for trials that did not report power calculations, several trials were clearly of inadequate size to achieve acceptable power at conventional levels of significance ($P<.05$).

Data Analysis

Analysis in 24 (44%) of the trials was confined to patients remaining "on randomized treatment"; 23 (42%) used an "intention to treat" approach, in which all subjects who are allocated to a particular intervention are included in the follow-up assessments irrespective of whether they had complied with the intervention requirements. Two trials used a combination of these methods, two used

Table 4. Quality Assessment of 55 Randomized Controlled Trials Published in Four US Primary Care Journals, 1987–1991

	Scores for Bias Control								
	Pharmacologic Trials			Nonpharmacologic Trials			Total Scores		
Potential Bias Variables	1 point	2 points	3 points	1 point	2 points	3 points	1 point	2 points	3 points
No. of trials controlling for bias before entry	13	4	3	28	4	3	41	8	6
No. of trials controlling for bias after entry*	5	10	4	11	6	18	16	16	22
No. of trials controlling for outcome assessment bias*	3	2	14	10	3	22	13	5	36

**In the one trial that described only the recruitment phase,[12] it was not possible to assess the quality of selection bias after entry or the bias in assessing the outcome.*

other methods, and four did not mention how the analysis was done.

In the 34 trials in which it was appropriate to address the issue of premature withdrawal of participants from the study, only 24 (71%) did so in the published reports. Similarly, in the 28 trials of pharmacological or nonpharmacological therapies in which it would have been appropriate to comment on the presence of any adverse effects, 19 (68%) raised the issue.

Assessment of Bias Control

There was a 96% observed agreement among the three reviewers regarding the extent to which bias was controlled in the trials included in the study. Of the three potential sources of bias examined, control of selection bias before entry was the most poorly controlled. Almost three fourths of the trials (n=40) made no attempt in their published reports to provide sufficient details to demonstrate that the randomization procedure had been done well and appropriately. A much greater proportion of trial reports provided information indicating that both selection bias after entry and bias in assessing the principal outcome had been controlled (Table 4). Selection bias after entry was controlled better in RCTs of nonpharmacological drug therapies or health service interventions than in trials that involved pharmacological agents. For the other two types of bias, there was no significant difference in the degree of control between trials of pharmacological agents and those of other types of interventions.

Organization and Management

Only six published reports included a statement that the study had been reviewed by an institutional review board; five other studies referred to participants providing informed consent before entry. In eight trials, the issue of field support staff was irrelevant. Of the remaining 48 studies, 11 (23%) involved research personnel who assisted in the practical organization and day-to-day operation; five (10%) reported using a practice nurse; and 31 (65%) did not mention whether any support staff were involved.

Discussion

The results of this study confirm that even though the absolute number of RCTs may be increasing, they still account for only a small proportion of articles published in primary care journals. This is disappointing, given the potential contribution of RCTs in providing useful evidence about the effectiveness of interventions frequently used in the discipline. It is important not to interpret this to mean that RCTs are either the most appropriate or the most desirable form of research methodology for many areas of family medicine research.

The results of a MEDLINE search conducted as part of the development of an international register of RCTs in primary care obtained across the same period suggest that only 26% of trials relevant to primary care are published in primary care or family medicine journals.[2] The remainder are found in the general medical literature and specialist nonprimary care literature.

Despite the small proportion of RCTs in primary care journals, the topics addressed using this research method are diverse, including a wide range of clinical and health service issues. Furthermore, a large percentage of RCTs reported in these four journals evaluated nonpharmacological therapies and health service interventions. Several of the trials reported in these journals have resulted in important new information about the discipline of family medicine.

It is difficult to make comparisons between journals because of the small number of RCTs during the study period and the differences in publication rates, which vary from monthly (*The Journal of Family Practice*) to quarterly (*Family Practice Research Journal*). Many of the findings in the present study, however, are similar to the findings in our survey of the *British Journal of General Practice* between 1958 and 1991.[3] The topic areas covered were generally similar to those reported in the devel-

oping international register of RCTs,[2] suggesting that a systematic bias in the topic areas of RCTs published in these four journals is unlikely.

Methodological characteristics of the trials were generally well reported. Analysis of these characteristics suggests that there is room for improvement in a few areas. For example, 49 (89%) of the trials failed to mention whether they had received prior ethical approval. Although the need for ethical approval is a relatively recent development, it is required not only for trials that involve pharmacological agents but also for those of nonpharmacological and health service interventions.

Many trials confined their data analyses to participants on randomized treatment only. Although this may be appropriate for trials examining the efficacy of a particular intervention among participants who comply with the treatment regimen, it limits the generalizability of the results to the "real world" of primary care and negates the main benefit of randomization. Including analyses of participants based on an "intention to treat" approach would reduce this problem in most trial reports. Furthermore, all participants who withdraw prematurely should be clearly accounted for wherever possible.

Another important deficit among these trials was the low number of papers reporting power calculations. It is important to pay attention to the sample size required to detect clinically relevant differences that may exist between various interventions (ie, adequate power), and to be reasonably certain that any difference found is not attributable to chance alone (ie, significance). Failure to do this diminishes the potential benefits arising from a well-conducted trial. Several trials in this series that failed to show a significant benefit for an intervention clearly had an inadequate sample size.

One of the frequent criticisms of general practice trials is a bias toward selecting healthy, cooperative patients who are most likely to benefit from participation.[13] It is impossible to confirm whether such a claim is valid without accurate details about the study population who were screened and found to be eligible for entry but not included. Such an "exclusion log" was presented in only one trial report in which the percentage of patients who were randomized represented only 20% of those eligible.[14]

The failure of many trials to achieve a high quality rating should be an incentive for researchers to specifically address each of the dimensions of potential bias in primary care trials. Of course, it is possible that appropriate steps to control bias were taken but were not documented in the published reports. Under these circumstances, it would be necessary to contact the authors of the reports to obtain further information about these methodological issues. We did not have the resources to do this.

It is encouraging that trials of nonpharmacological therapies and health service interventions control bias more effectively than do trials of pharmacological therapies, since it demonstrates that it is possible to achieve high-quality RCTs in these areas as well. The relatively poor control of bias seen in trials of pharmacological agents is surprising because control for selection bias at entry is easily ensured in such studies by having an independent source, such as a pharmacy, prepare identical preparations of the agents involved and arrange for the randomization procedure. Furthermore, since many drug trials are placebo-controlled, using a blind assessment of trial outcome measures should be easy.

The findings reported in this study are not intended as a criticism of either the journals or the discipline of family medicine. Reviews of methodological characteristics and quality assessment in other disciplines have reported substantial variation in the standards found in published reports.[9] It is encouraging to find so many excellent examples of imaginative and effective solutions to many of the problems associated with the design and conduct of clinical trials in primary care. These solutions should serve as an incentive to those undertaking clinical trials in the future.

For example, in a trial that compared topical antibiotic ointments, wound protectants, and antiseptics for treatment of human blister wounds, patients were inoculated so as to produce three small artificial wounds on each forearm. The wounds were then randomized to receive one of the various interventions or remain untreated as a control. In this way, each participant acted as his or her own control.[15]

In another trial comparing two different educational strategies to promote the use of seat belts among school children, outcome assessment was handled in an unusual way. A research assistant, who acted as the observer and was unaware of which students had been allocated to which intervention group, was situated on the roof of the school building with a pair of binoculars to observe and record whether students traveling home from school by car wore their seat belts.[16]

The value of the randomized controlled trials published in these journals ultimately depends on the extent to which they appropriately influence practice in the light of other available evidence. A new international undertaking, the Cochrane Collaboration in Primary Health Care,[17] is attempting to systematically identify, assemble, collate, and synthesize information from all randomized controlled trials (published and unpublished) relevant to primary care. The value of this process and the information it provides will to a large extent reflect the quality of the individual trials included. Investigators, journal editors, and reviewers all have a responsibility to ensure that high standards are set and maintained for RCTs in this

discipline. The criteria used in this study to assess the trial reports may serve as a useful starting point for investigators in their preparation of future clinical trial reports.

Acknowledgments

The authors wish to thank Dr Iain Chalmers, director of the UK Cochrane Centre, Oxford, for advice and encouragement; the Imperial Cancer Research Fund Library for assisting with the MEDLINE search; and S. Nightingale, for retrieving and photocopying relevant articles.

Funding for Dr Silagy was provided by the Sir Robert Menzies Memorial Trust in London.

References

1. Sackett DL, Haynes RB, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. Boston, Mass: Little, Brown, 1985.
2. Silagy CA, Jewell MD. 39 years of randomised controlled trials in the *British Journal of General Practice*: an analysis. Br J Gen Pract 1994; 44:359–63.
3. Silagy CA. Developing a register of randomized controlled trials in primary care. BMJ 1993; 306:897–900.
4. Gruppo Di Studio Sulla Pressione Arteriosa Nell'Anziano. Feasibility of a large prospective study in general practice: an Italian experience. BMJ 1987; 294:157–60.
5. Franks P. Clinical trials. Fam Med 1988; 20:443–8.
6. Hvidberg EF, Pedersen PA. Therapeutic trials in general preactice, an unmet need. Dan Med Bull 1990; 37:84–6.
7. Aberg H. Multicentre trials run general practice. Scand J Prim Health Care 1989; 7:129–30.
8. Chalmers TC, Smith H Jr, Blackburn B, et al. A method for assessing the quality of a randomized controlled trial. Controlled Clin Trials 1981; 2:31–49.
9. WONCAGP (World Organization of National Colleges and Academies of General Practice). An international classification of the health problems of primary care (ICHPPC-2). 3rd ed. Oxford, England: Oxford University Press, 1983.
10. Chalmers I, Enkin M, Keirse MJNC, eds. Effective care in pregnancy and childbirth. Oxford, England: Oxford University Press, 1985.
11. Gardner MJ, Gardner SB, Winter PD. Confidence interval analysis, version 1.1. London, BMJ, 1991.
12. Clinical Experience Network (CEN). Low high-density lipoprotein cholesterol and other coronary heart disease risk factors in patients with total cholesterol levels greater than 5.17 mmol/L (200 mg/dL) in family practice. J Am Board Fam Pract 1991; 4:285–97.
13. Bope ET, Bates CM. Drug trial research in family practice. Fam Pract Res J 1988; 7:187–9.
14. Dunlay J, Reinhardt R, Roi LD. A placebo-controlled, double-blind trial of erythromycin in adults with acute bronchitis. J Fam Pract 1987; 25:137–41.
15. Leyden JL, Bartelt NM. Comparison of topical antibiotic ointments, a wound protectant, and antiseptics for the treatment of human blister wounds contaminated with *Staphylococcus aureus*. J Fam Pract 1987; 24:601–4.
16. Gordon P. Evaluation of an intervention to increase seat belt use. Fam Med 1989; 21:458–60.
17. Silagy CA, Lancaster T. The Cochrane Collaboration in Primary Health Care. Fam Pract 1993; 10:364–5.