J Clin Epidemiol Vol. 50, No. 2, pp. 137–146, 1997

0895-4356/97/$17.00
PII S0895-4356(96)00338-1

Are Community Health Interventions Evaluated Appropriately? A Review of Six Journals

Pamela J. Smith,[1] *Michael E. K. Moffatt*[1,2,*]
Shirley C. Gelskey,[1,3] *Shauna Hudson,*[1] *and Kelly Kaita*[4]

[1]Department of Community Health Sciences, [2]Department of Pediatrics, [3]Department of Dental Diagnostics and Surgical Sciences, and [4]Department of Internal Medicine, Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0W3

ABSTRACT. *Objectives:* To determine if Randomized Controlled Trial (RCT) methodology was used appropriately in community health, we: (1) determined the proportion of non-randomized studies that should have been RCTs, and (2) assessed the quality of the RCTs. *Methods:* The 1992 issues of six community health journals were manually searched. Intervention studies were analyzed. Studies that did not use randomization were analyzed for feasibility and practicality of RCT methods; RCTs were analyzed for quality using a checklist. RCTs were compared with community health RCTs from *The New England Journal of Medicine*. The proportion of studies meeting each criterion was determined. *Results:* Fourteen percent of 603 studies were interventions and 4% were RCTs. Of those not using randomization, 42% should have. Mean RCT scores were significantly lower for the community health journals than for *The New England Journal of Medicine*. Many criteria important to quality scored poorly. *Conclusions:* RCTs are underused and lack methodologic rigor in community health. Conclusions regarding the effectiveness of interventions are therefore suspect. *Copyright © 1997 Elsevier Science Inc.* J CLIN EPIDEMIOL 50;2:137–146, 1997.

KEY WORDS. Interventions, randomized controlled trials, literature search, methodology, quality assessment

INTRODUCTION

> The health sector is subject to a severe inflation with the output rising much less than would be expected from the input . . . the inflation could be controlled by science, in particular by the wide use of randomized controlled trials [1].

More than two decades ago, Cochrane [1] recognized that well-designed intervention studies were necessary to establish the efficacy and effectiveness of health promotion and preventive maneuvers. The inflation described by Cochrane is even more of a concern today. Due to their wide application, costs of community health interventions can be monumental. With increasing pressure to use resources efficiently, "evidence-based" programs and practices are gaining emphasis in other areas of medicine [2–6].

Recently, the Cochrane Collaboration was formed to promote evidence-based practice by coordinating international efforts to prepare and disseminate systematic reviews of Randomized Controlled Trials (RCTs) [7]. RCT research methodology is regarded as the most scientifically rigorous means to evaluate interventions [8]. Despite claims of its demise, as well as the lackluster results of huge field trials, at "unconscionable expense" [8], we believe that RCTs have a role to play in community health research. However, it was our impression that RCT methods were underused and lacked methodologic rigor in community health research. By reviewing all original research published in one year in six community health journals, the objectives of this study were to: (1) determine the number of RCTs and the proportion of the remaining intervention studies that should, in our opinion, have used RCT methodology, (2) assess the internal validity and overall quality of the RCTs in comparison to community health RCTs in *The New England Journal of Medicine*, and (3) compare the manual RCT search with a MEDLINE search.

METHODS

The last complete year, 1992, of six locally available journals selected for their general community health focus (*Am J Epidemiol, Am J Public Health, Am J Prev Med, Can J Public Health, J Public Health Med,* and *Public Health Rep*) were manually searched for original research. Non-intervention studies were counted and excluded. All studies assessing ef-

*Address for correspondence: Dr. Michael Moffatt, Department of Community Health Sciences, University of Manitoba, S100-750 Bannatyne Avenue, Winnipeg, Manitoba, Canada R3E 0W3.
Accepted for publication on 11 September 1996.

fects of an intervention (program, policy, or treatment) were categorized as community health interventions and as either non-randomized intervention studies (NRISs) or RCTs based on allocation method [9].

Community health RCTs, as opposed to other types (i.e., trials of treatment regimens for individuals), were identified as those RCTs seeking results applicable by public health professionals on a preventive basis. Community health RCTs from the 1992 issues of *The New England Journal of Medicine* were scored using the same criteria. These RCTs were retrieved by three independent manual searches (PS, MM, SG) and collective agreement was reached on inclusion criteria.

The Non-Randomized Intervention Studies (NRISs)

NRISs were categorized as those that (1) could not have been an RCT, (2) could but should not have been an RCT, and (3) could and should have been an RCT. Each NRIS was assessed by a single author for stated or inferred rationale for not using randomization and for the feasibility and practicality of an RCT approach. These evaluations were based on the reviewer's interpretation of study design and context of research (pilot study or part of larger study); rationale; previous evidence of efficacy; ease with which randomization could have been used; and a value judgement as to the importance of the research question relative to potential cost increases associated with RCT methodology. During roundtable meetings, reviewers' opinions were discussed and each NRIS was categorized based on a consensus decision.

The Randomized Controlled Trials

RCTs were assessed for adherence to criteria important to quality of design, conduct, analysis, and report, with emphasis on internal validity. A dichotomous checklist (Table 1), similar to one used by Orr [10], was adapted from the Chalmers scale [11]. The Chalmers scale emphasizes features that may be difficult to incorporate into community health RCTs (i.e., quadruple blinding and testing of blinding) whereas our checklist was developed to identify problem areas. None of the items are specific to community health.

Apart from sampling error, differences between groups can be attributed to the hypothesis under investigation only if the study is internally valid [9]. Of 35 items, 26 pertain to internal validity. The remaining nine criteria concern external validity (generalizability beyond study population) and quality of reporting objectives, conclusions, and ethical consent. The checklist was pretested for ease and comparability of use, but not for inter-rater reliability, by two authors using three 1991 RCTs. Critical review required approximately 20 minutes per RCT.

Each of the 33 RCTs was scored independently by at least three authors (SG, MM, PS; KK also scored 18 studies). The items were scored 0 or 1, except in rare instances (12/1155 scores) when 0.5 was scored for criteria addressed but not demonstrated. Scores were compared and discussed item by item for each study during conferences. Consensus was not reached for 22 (2%) of the 1155 scores: 10 times for conclusions supported by results; 3 times for both clear objectives and objective or blindly assessed outcome; twice for recruitment methods; and once for each of items 8, 19, 22, and 26 (see Table 1). Errors and omissions were corrected before calculating final scores and differences in opinion were resolved by averaging scores. Non-applicable criteria were not included in denominators. Although most criteria do not require supplemental instructions, to meet specific criteria: blinding was explicitly reported or inevitable from study design; a method of randomization that concealed assignment until after allocation was reported; and an acceptable method of handling withdrawals was reported if withdrawals were greater than 5%.

Disguising the source of the articles would have required retyping, thus no attempt was made to blind the reviewers. To prevent all articles from a single journal being evaluated in sequence, the studies were assessed in a pre-specified systematic order. Prior publications were retrieved when necessary to score criteria.

Statistical Analysis

Mean scores for internal validity and overall quality of RCTs were calculated for the six community health journals combined and for *The New England Journal of Medicine* alone. A Students' *t*-test was used to compare these mean scores. The frequency distribution of scores for each criterion was determined.

The following descriptive data were collected for each RCT to determine if any were statistically associated with quality scores: study duration; country; unit of randomization; funding sources; multiple/single-center study status; statistical consultation acknowledgment; type of study (reviewers' opinion: efficacy/effectiveness); subject area; authors' credentials (i.e., Ph.D. M.D.); and authors' affiliations (i.e., University, Research Institute).

The Medline Search

Subsequent to the manual search, an attempt was made to identify the RCTs in the six community health journals using only MEDLINE [12]. Several strategies were employed using terms selected from MEDLINE Index. MEDLINE entries for missed RCTs were examined for terms to improve the MEDLINE retrieval rate.

RESULTS

The 603 original research reports in the six community health journals are classified by study type in Fig. 1. Eighty-two (14%) are intervention studies (57 NRISs [13–69], 25 RCTs [70–94]). The distribution of intervention studies is

TABLE 1. Criteria for evaluating RCT quality, and percentage of studies meeting each criterion

%	Criteria	%	Criteria
	OBJECTIVES		OUTCOME
91	1. Objectives clearly stated?	97[a]	19. Outcome measures defined?
	TARGET POPULATION	58[a]	20. Outcome either objective or assessed blindly?
94	2. Target population clear?	70[a]	21. Validity/reliability of outcome measure assessed/known?
	SAMPLE		STUDY CONFOUNDERS
94[a]	3. Inclusion criteria defined?	13[a]	22. Contamination assessed?
33	4. Sample representative of target population?	70[a]	23. Compliance assessed?
12[a]	5. Sample calculated prior to study?	41[a]	24. Side effects measured?
70	6. Recruitment methods described?	30[a]	25. Differences between withdrawals and those completing trial assessed?
93[a]	7. Exclusion criteria defined?	37[a]	26. Acceptable method of handling withdrawals given?_____
42	8. Number of eligible and, if applicable, ineligible subjects and refusals?	38[a]	27. Total number of withdrawals acceptable (<15%)?_____
10	9. Differences between refusals and participants assessed?		STATISTICAL ANALYSIS
	INTERVENTION	91[a]	28. Appropriate statistical tests used?
100[a]	10. Intervention described?	52[a]	29. Adjustment for baseline characteristic differences?
21[a]	11. Co-intervention controlled for?	11[a]	30. Level of Type II error stated for negative results?
97[a]	12. Control and intervention maneuvers appropriate?	88[a]	31. p Value given?
	RANDOMIZATION	48[a]	32. Confidence limits given?
9[a]	13. Acceptable method reported?		CONCLUSIONS
58[a]	14. Intergroup balance of baseline characteristics assessed?	48	33. Conclusions supported by results?
	BLINDING		ETHICS
53[a]	15. Subjects blinded?	74	34. Informed consent addressed?
15[a]	16. Study team blinded?	18	35. Approval of ethics review board?
0[a]	17. Blinding of subjects tested?		
0[a]	18. Blinding of study team tested?		

Source: Adapted from Chalmers *et al.* 1981 [11].
[a]Criteria pertaining to internal validity.

presented by journal in Fig. 2 with the NRISs divided according to our recommendations for use of RCT methodology.

The Non-Randomized Intervention Studies (NRISs)

The NRISs are further categorized by recommended methodology and stated or inferred reasons for not using randomization (Fig. 1). Of the 57 NRISs, 14 (25%) [13–26] could not have been RCTs; and 19 (33%) [27–45] could have been but should not have been RCTs. However, 24 NRISs (42%) [46–69] could and should have been RCTs to best address the research question. Thus, although 25 (30%) of the 82 intervention studies were RCTs, in our opinion 49 (60%) should have been RCTs.

The Randomized Controlled Trials

In 1992, 25 RCTs (30% of intervention studies; 4% of all studies) were published in the six community health journals and 8 community health RCTs were published in *The New England Journal of Medicine* [95–102]. Of the 33 RCTs, one used a crossover design [71] and one used a factorial design [81]. The individual was the unit of randomization in 24 RCTs, the group in 8 [82–84,88,91,93,100,102], and both the individual and group in one [76]. The topics addressed were nutrition [71,81,86,88,90,91,97,99–101], smoking [73,75,77,78,82,85,89], AIDS/HIV [70,74,76, 84,94,95], mammography [79,92,93], seniors' programs [72,80,102], health-care delivery [83,87], and prenatal programs [96,98].

Internal validity and overall scores for individual RCTs ranged from 26–79% and 27–77%, respectively. Mean RCT scores were significantly lower for the community health journals combined ($n = 25$) than for *The New England Journal of Medicine* ($n = 8$) ($p < .05$). The mean scores for internal validity and overall quality, respectively, were 48.3% (range 26.1–79.2, SD = 11.9) and 50.4% (range 27.1–75.8, SD = 11.8) for the community health journals combined compared with 65.2% (range 50.0–75.0, SD = 7.9) and 64.7% (range 48.5–77.1, SD = 9.8) for *The New England Journal of Medicine*.

The proportion of RCTs that met each criterion is given in Table 1. The most poorly handled internal validity criteria were testing of blinding of both subjects (0%) and study team (0%); providing method of randomization (9%); providing level of type II error for negative results (11%); calculating sample size prior to study (12%); assessing contamination (13%); and blinding of study team (15%).

In this review, the three criteria most important to generalizability were poorly met: differences between refusers and participants assessed (10%); study sample representative of

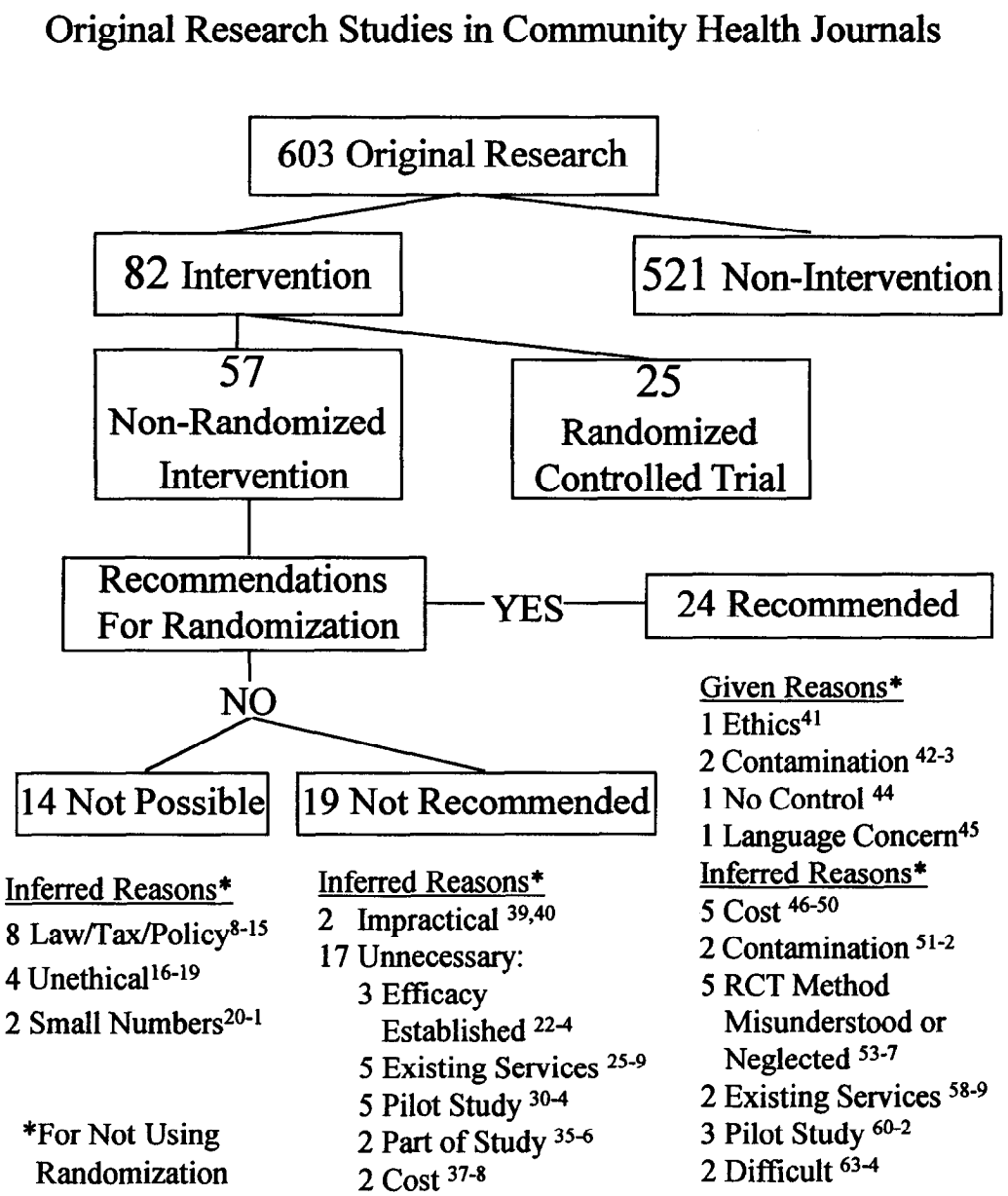

FIGURE 1. Results of the manual literature search and classification by methods and recommendations for using RCT methods. For the non-randomized intervention studies, the given or inferred reasons for not having been conducted as RCTs are provided (superscript numbers).

the target population (33%); and numbers of eligible subjects, ineligible subjects and refusals reported (42%).

Statistical analysis using the descriptive data was not possible due to small numbers and lack of variability. It was possible to identify statistical expertise among the authors' credentials or the acknowledgments for only six RCTs and all but two RCTs were single-center studies. Three were authored by M.D.s only, nine by PhDs only, fourteen by both, and four by people with other credentials. Ten RCTs acknowledged more than one source of funding. Most commonly, research agencies (20), government (13), or both (6) provided funds.

The MEDLINE Search

The MEDLINE search using the index term RANDOMIZED-CONTROLLED-TRIAL retrieved 15 of the 25 RCTs found in the manual search of the six community health journals. The search included a further 10 irrelevant citations, encompassing meta-analyses, methodology papers, and descriptions without results.

Addition of the Index term RANDOM-ALLOCATION netted one more RCT. Review of the complete MEDLINE entry for all missed RCTs revealed several abstracts containing "randomly assigned" or "assigned randomly." Inclu-

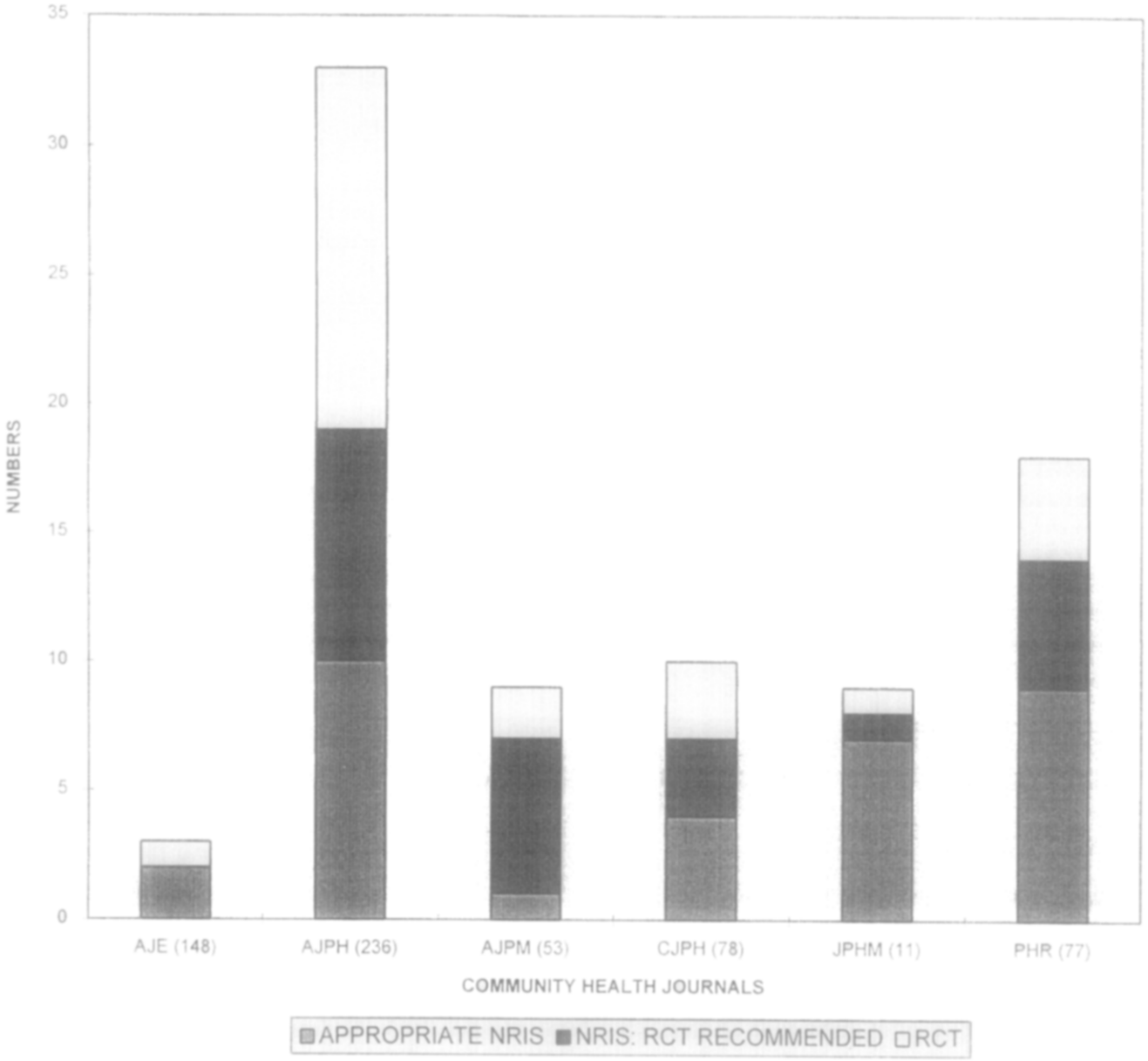

FIGURE 2. The number of interventions by journal: RCTs (*white*) and NRISs (*all shaded portions*), with the NRISs divided (into appropriate NRIS and RCT recommended) according to our recommendations. The total number of original articles in each journal is under the bar for each journal in parentheses.

sion of this phrase, instead of RANDOM-ALLOCATION, retrieved 21 of 25 RCTs and 10 irrelevant references. There were no clues in the MEDLINE entry to indicate that three of the remaining four RCTs were, in fact, RCTs. The abstract of the fourth RCT described the randomization clearly, but the only term useful for a MEDLINE search was "random" which also recovers immense numbers of studies using random sampling and random-digit dialing.

DISCUSSION

Because unknown and unintentional bias can render studies invalid "observational evidence is never very satisfactory" [1]. Because randomization is the best way to control bias [103,104], RCTs are generally regarded as the "gold standard" for evaluation of the effectiveness of interventions [105,106].

The Non-Randomized Intervention Studies (NRISs)

For many reasons, including cost, ethics, rarity of outcome, and an inability to randomize the intervention, RCTs can never make up 100% of the intervention studies in community health research. Nonexperimental research has been responsible for some extraordinary health achievements and its role must not be minimized [105,107].

Randomization is not a remedy for poor design, execution, or analysis [103]. Good quality observational studies may produce results that are more valid than poorly done RCTs. It is also important not to design community health RCTs that are too ideal or unrealistic. Efficacy established under extreme conditions of control and selectivity may not translate into effectiveness when the intervention is launched on a wide-scale basis [108].

In this review, we considered it appropriate that 33 NRISs were not RCTs for ethical and practical reasons: 15 studies could not have been randomized and 18 could have but should not have been RCTs. This latter category included studies for which additional costs of RCT methods would not be justified: pilots, audits, and interventions with established efficacy.

For at least 5 [58–62] of the 24 NRISs which should have been RCTs, a lack of awareness or misunderstanding of RCT methods appears to account for the failure to use randomization. Only 5 of the 43 NRISs that could have been RCTs reported any rationale for not conducting an RCT [46–50]. Researchers may not consider RCTs to be part of community health research methods, and there may be hesitancy to use control groups. If true, this is indefensible. Unless an intervention has been shown to be effective it would be fiscally irresponsible and potentially harmful to adopt it.

Potential confounders are less amenable to control in community health research than in drug and animal studies. Although blinding is an important RCT method, it cannot always be used in community health interventions, and individuals cannot always be randomized. However, with effort, blinding can often be used, and randomization of groups or communities may be possible. Group randomization may facilitate blinding of both providers and participants to the evaluative nature of the study. For example, an educational program could be tested in randomly selected communities.

The Randomized Controlled Trials

We found considerable variation in quality among RCTs. Descriptions of the interventions, outcome measures, target populations, inclusion and exclusion criteria, and objectives, as well as the appropriate use of control and intervention maneuvers were criteria that were satisfactorily handled by all 33 RCTs. Some problems identified in the RCTs in this review may result from reporting deficiencies as opposed to design flaws [109]. Such ambiguity is unacceptable; the reader cannot assess study validity.

Blinding criteria scored poorly, but blinding is problematic in community health research. Other faults are potentially more serious. Blinding after allocation is not always possible but concealment of randomization up to the point of allocation is *always* possible and is crucial for successful randomization [110].

Methods of randomization vary in their ability to conceal assignment and avoid bias. Sealed envelopes are less subject to bias than coin tossing [111]. Reporting an acceptable method of randomization is rare [110,112,113]; only 9% of the RCTs in this review did so. For at least two RCTs, allocation was probably not random but systematic [83,92]. Studies reporting blocked designs or stratified random samples may have used an acceptable method of generating random assignments but the methods must be explicitly reported [113]. In one trial [88] the randomization process was clearly not concealed and thus subject to bias [110,114]. Inadequate or unclear allocation has been associated with larger estimates of differences between intervention groups [113–116].

After randomization the similarity of baseline characteristics must be assessed. Although randomization will allocate without bias, it will not necessarily generate equivalent groups [117]. Only 58% of the RCTs made this assessment. Of the studies without demonstrated similarity of baseline characteristics, only 52% used statistical methods to adjust for potential differences.

Results must be analyzed in the group to which the subjects were randomly allocated, whether or not they completed or even received that intervention. Failure to use this intention-to-treat analysis can produce misleading results [118]. One of the studies [82] invalidated the randomization process by the reassignment of subjects for analysis according to their compliance.

With group randomization, the responses of individuals within groups are not statistically independent. Variation within groups is likely to be less than between groups. Clustering must be considered in sample size calculations and statistical analyses. Failure to do so leads to underestimated sample sizes and narrower confidence intervals that can generate statistically significant results spuriously [119]. Only one of nine studies using group randomization reported accounting for clustering when calculating sample size prior to the study [100], and another discussed power implications of clustering [102]. Seven studies took clustering into account in the analysis [82–84,88,91,93,100,102].

Only 3 (11%) of 27 studies reporting negative results reported the level of type II error. The power of a study to detect a significant difference has serious implications for the way in which a negative result should be interpreted [111,120]. Reporting that two maneuvers are equivalent is misleading if sufficient statistical power to detect a difference does not exist. The power of a study must be considered before the study begins rather than after the analysis; it will determine the sample size necessary to make the study meaningful [121,122]. Only 12% of RCTs reported prestudy calculation of sample size.

In their guidelines for assessing RCT quality, Chalmers *et al.* [10] recommend careful scrutiny of trials with more than 15% withdrawals. While some degree of loss to follow-up is usually inevitable, the internal validity of the study may not be seriously compromised if the number of with-

drawals is low, equivalent in all study groups, and the baseline characteristics of those who completed the study do not differ significantly from those lost to follow-up. Fifteen RCTs (45%) were analyzed excluding withdrawals [70–74,76,77,79,86,88–90,93,94,97]. However, 59% of the RCTs in this review had more than 15% withdrawals and 70% of the 27 studies that should have made comparisons on completion status did not. Ten RCTs appropriately included the withdrawals in the analysis as "failures" in the group to which they were assigned [75,78,80,81,84,85,91, 96,98,100]. One non-dichotomous study conservatively used a measure of twice the rate of adverse events in withdrawals than in subjects who completed the trial [95]. For three studies, the method of handling withdrawals was unknown [87,92,102].

The significantly higher mean scores for *The New England Journal of Medicine* RCTs might indicate that the better studies were submitted to and accepted by this oft-cited journal (Science Citation Index "impact score" for *The New England Journal of Medicine* was 23.223 while the highest reviewed community health journal scored 3.190 [123]). However, this does not account for the mediocre scores attained by all the RCTs, especially since most had authors affiliated with University departments of epidemiology and all journals were peer-reviewed.

Notably, only 48% of the 33 RCTs reported conclusions we felt were supported by their data. Our findings indicate significant deficiencies in the design, conduct, and reporting of RCTs by community health researchers, and reflect a lack of well-specified standards for publication of RCTs. Simply by virtue of randomization, regardless of quality, RCTs are less prone to publication bias [124]. Because RCT methodology confers scientific credibility and carries great weight in the medical community, it is especially important to maintain high standards of acceptability for publication [105,106]. To this end, guidelines for structured reporting have been published [125,126]. Authors and editors could benefit from the use of checklists, while those attempting meta-analyses should consult directly with authors and use a scale to rate the likelihood of freedom from bias [127].

The MEDLINE Search

Results of the comparison of MEDLINE and manual searches will be of interest to those conducting systematic reviews or meta-analysis of RCTs from community health journals. The manual search was assumed to have a sensitivity of 100%; methods sections of all articles were reviewed and all intervention studies were selected for careful perusal. The specificity of the manual search was 78% (seven papers were not RCTs). The best MEDLINE search strategy used both the MEDLINE Index term RANDOMIZED-CONTROLLED-TRIAL and the phrase "randomly assigned" and had a sensitivity of 84% and specificity of 68%. This comparison of the results of the two search procedures highlights the need to conduct manual searches to ensure complete retrieval of RCTs. Ideally it should not be this difficult to determine if a study employed RCT methods. The routine use of structured abstracts [128,129] might help to solve this problem, as would restricting use of the term RANDOMIZED-CONTROLLED-TRIAL in the MEDLINE PUBLICATION-TYPE field to only those studies *using* RCT methodology.

CONCLUSIONS

The results of this review suggest that many interventions in community health research are not evaluated appropriately. RCTs are underused in community health research and that RCTs published in 1992 in six journals of community health generally lacked methodologic rigor. The use of RCT methodology to evaluate interventions in community health research should be encouraged to a much greater extent, along with greater attention to quality of design, conduct, analysis, and report.

This research was made possible in part by Health Canada through a National Health Research and Development Program research training fellowship to Pamela J. Smith.

References

1. Cochrane AL. **Effectiveness and Efficiency: Random Reflections on Health Services**. London: Nuffield Provincial Hospitals Trust; 1972.
2. Evidence-Based Care Resource Group. Evidence-based care: 1. Setting priorities: How important is this problem? **Can Med Assoc J** 1994; 150: 1249–1254.
3. Evidence-Based Care Resource Group. Evidence-based care: 2. Setting guidelines: How should we manage this problem? **Can Med Assoc J** 1994; 150: 1417–1422.
4. Evidence-Based Care Resource Group. Evidence-based care: 3. Measuring performance: How are we managing this problem? **Can Med Assoc J** 1994; 150: 1575–1579.
5. Rafuse J. Evidence-based medicine means MDs must develop new skills, attitudes, CMA conference told. **Can Med Assoc J** 1994; 150: 1479–1481.
6. Evidence-Based Medicine Working Group. Evidence-based medicine: A new approach to teaching the practice of medicine. **JAMA** 1992; 268: 2418–2425.
7. Robinson A. Research, practice and the Cochrane Collaboration. **Can Med Assoc J** 1995; 152: 883–889.
8. Herman J. The demise of the randomized controlled trial. **J Clin Epidemiol** 1995; 48: 985–988.
9. Last JM, Ed. **A Dictionary of Epidemiology**. 2nd ed. Toronto: Oxford University Press; 1988.
10. Orr PH, Scherer K, Macdonald A, Moffatt MEK. Randomized placebo-controlled trials of antibiotics for acute bronchitis: A critical review of the literature. **J Farn Pract** 1993; 36: 507–512.
11. Chalmers TC, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomized control trial. **Controlled Clin Trials** 1981; 2: 31–49.
12. **MEDLINE** (R) SilverPlatter (R) 3.11. Medline Express 1989–1993.

13. Williams AF, Wells JK, Lund AK, Teed NJ. Use of seatbelts in cars with automatic belts. **Public Health Rep** 1992; 107: 182–188.
14. Hinds MW. Impact of a local ordinance banning tobacco sales to minors. **Public Health Rep** 1992; 107: 355–358.
15. Paulozzl LJ, Spengler RF, Gower M-A. An evaluation of the Vermont worksite smoking law. **Public Health Rep** 1992; 107: 724–726.
16. Taylor DH, Ricketts TC, Berman JL, Kolimaga JT. One state's response to the malpractice insurance crisis: North Carolina's Rural Obstetrical Care Incentive Program. **Public Health Rep** 1992; 107: 523–529.
17. DiFranza JR, Brown LJ. The Tobacco Institute's "It's the Law" campaign: Has it halted illegal sales of tobacco to children? **Am J Public Health** 1992; 82: 1271–1273.
18. Forster JL, Hourigan ME, Kelder S. Locking devices on cigarette vending machines: Evaluation of a city ordinance. **Am J Public Health** 1992; 82: 1217–1219.
19. Peterson DE, Zeger SL, Remington PL, Anderson HA. The effect of state cigarette tax increases on cigarette sales, 1955 to 1988. **Am J Public Health** 1992; 82: 94–96.
20. Flewelling RL, Kenney E, Elder JP, Pierce J, Johnson M, Bal DG. First-year impact of the 1989 California cigarette tax increase on cigarette consumption. **Am J Public Health** 1992; 82: 867–869.
21. Poulin C, Gyorkos TW, Joseph L. The hepatitis B immunization strategy in Nova Scotia: Is it effective? **Can J Public Health** 1992; 83: 286–289.
22. Robertson SE, Markowitz LE, Berry DA, Dini EF, Orenstein WA. A million dollar measles outbreak: Epidemiology, risk factors, and a selective revaccination strategy. **Public Health Rep** 1992; 107: 24–31.
23. Osberg JS, DiScala C. Morbidity among pediatric motor vehicle crash victims: The effectiveness of seat belts. **Am J Public Health** 1992; 82: 422–425.
24. Esrey SA, Habicht J-P, Casella G. The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. **Am J Epidemiol** 1992; 135: 659–666.
25. Morin R, Chapdelaine A, Schaefer C. La promotion du coussin gonflable: Evaluation d'une intervention de sante publique aupres de candidats aux elections federales. **Can J Public Health** 1992; 83: 234–236.
26. Ahmann E, Wulff L, Meny RG. Home apnea monitoring and disruptions in family life: A multidimensional controlled study. **Am J Public Health** 1992; 82: 719–722.
27. Foster DA, Talsma A, Furumoto-Dawson A, Ohmit SE, Marguilies JR, Arden NH, Monto AS. Influenza vaccine effectiveness in preventing hospitalization for pneumonia in the elderly. **Am J Epidemiol** 1992; 136: 296–307.
28. Higginson IJ, Wade AM, McCarthy M. Effectiveness of two palliative support teams. **J Public Health Med** 1992; 14: 50–56.
29. Manea SJ, Iohp K. The hepatitis B immunization campaign for children in the Federated States of Micronesia. **Public Health Rep** 1992; 107: 556–561.
30. O'Mara LM, Isaacs S, Chambers LW. Followup of participants in a preschool hearing screening program in child care centres. **Can J Public Health** 1992; 83: 375–378.
31. Ahmed S, Hicks NR, Stanwell-Smith R. Policy and practice—An audit of neonatal BCG Immunization in Avon. **J Public Health Med** 1992; 14: 389–392.
32. Schramm WF. Weighing costs and benefits of adequate prenatal care for 12,023 births in Missouri's Medicaid Program, 1988. **Public Health Rep** 1992; 107: 647–652.
33. Hartgers C, van Ameijden EJC, van den Hoek JAR, Coutinho RA. Needle sharing and participation in the Amsterdam Syringe Exchange Program among HIV-seronegative injecting drug users. **Public Health Rep** 1992; 107: 675–681.
34. Hartgers C, van den Hoek A, Krijnen P, Coutinho RA. HIV prevalence and risk behavior among injecting drug users who participate in "Low-Threshold" methadone programs in Amsterdam. **Am J Public Health** 1992; 82: 547–551.
35. Roland M, Bewley B, Boneline: Evaluation of an initiative to improve communication between specialists and general practitioners. **J Public Health Med** 1992; 14: 307–309.
36. Ebrahim S, Williams J. Assessing the effects of a health promotion programme for elderly people. **J Public Health Med** 1992; 14: 199–205.
37. Payne JN. The introduction of a computerized system for notification and improved analysis of infectious diseases in Sheffield. **J Public Health Med** 1992; 14: 62–67.
38. Flynn BS, Worden JK, Secker-Walker RH, Badger GJ, Geller BM, Constanza MC. Prevention of cigarette smoking through mass media intervention and school programs. **Am J Public Health** 1992; 82: 827–834.
39. Cotugna N, Vickery CE. Development and supermarket field testing of videotaped nutrition messages for cancer risk reduction. **Public Health Rep** 1992; 107: 691–694.
40. Godel JC, Pabst HF, Hodges PE, Johnson KE. Iron status and pregnancy in a Northern Canadian population: Relationship to diet and iron supplementation. **Can J Public Health** 1992; 83: 339–343.
41. Perry CL, Kelder SH, Murray DM, Klepp K-I. Community-wide smoking prevention: Long-term outcomes of the Minnesota Heart Health program and the class of 1989 study. **Am J Public Health** 1992; 82: 1210–1216.
42. McIlvain HE, McKinney ME, Thompson AV, Todd GL. Application of the MRFIT smoking cessation program to a healthy, mixed-sex sample. **Am J Prev Med** 1992; 8: 165–170.
43. Houghton A, Egan S, Archinal G, Bradley O, Azam N. Selective medical examination at school entry: should we do it, and if so how? **J Public Health Med** 1992; 14: 111–116.
44. Bull AD, Silcocks PB, Start RD, Kennedy A. General practitioners, skin lesions and the new contract. **J Public Health Med** 1992; 14: 300–306.
45. Andrews EB, Joseph MC, Magenhelm MJ, Tilson HH, Doi PA, Schultz MW. Postmarketing surveillance study of permethrin creme rinse. **Am J Public Health** 1992; 82: 857–861.
46. Scheffer RM, Feuchtbaum LB, Phibbs CS. Prevention: The Cost-effectiveness of the California diabetes and pregnancy program. **Am J Public Health** 1992; 82: 168–175.
47. Kawatu D, MacLachlan RA, Lees REM. Intestinal parasite control in an institution for developmentally handicapped adults. **Can J Public Health** 1992; 83: 295–297.
48. Kellerman ST, Felts WM, Chenier TC. The impact on factory workers of health risk appraisal and counselling in health promotion. **Am J Prev Med** 1992; 8: 37–42.
49. Mayer JA, Kossman MK, Miller LC, Crooks CE, Slymen DJ, Lee CD, Jr. Evaluation of a media-based mammography program. **Am J Prev Med** 1992; 8: 23–29.
50. Parsons L, Day S. Improving obstetric outcomes in ethnic minorities: An evaluation of health advocacy in Hackney. **J Public Health Med** 1992; 14: 183–191.
51. Birdsell JM, Campbell HS, McGregor SE, Hill GB. Steve Fonyo cancer prevention program: Description of an intervention. **Can J Public Health** 1992; 83: 237–239.
52. Buescher PA, Ward NI. A comparison of low birth weight among Medicaid patients of public health departments and other providers of prenatal care in North Carolina and Kentucky. **Public Health Rep** 1992; 107: 54–59.

53. Helvie CO, Alexy BB. Using after-shelter case management to improve outcomes for families with children. **Public Health Rep** 1992; 107: 585–588.
54. Kelly JA, St. Lawrence JS, Stevenson Y, Hauth AC, Kalichman SC, Diaz YE, Brasfield TL, Koob JJ, Morgan MC. Community AIDS/HIV risk reduction: The effects of endorsements by popular people in three cities. **Am J Public Health** 1992; 82: 1483–1489.
55. Calysn DA, Saxon AJ, Stanton V. Risk reduction in sexual behaviour: A condom giveaway program in a drug abuse treatment clinic. **Am J Public Health** 1992; 82: 1536–1538.
56. Greenlee P, Castle CH, Woolley FR. Successful modification of medical students' cardiovascular risk factors. **Am J Prev Med** 1992; 8: 43–52.
57. Baier CA, Grodzin CJ, Port JD, Leksas L, Tancredi DJ. Coronary risk factor behaviour change in hospital personnel following a screening program. **Am J Prev Med** 1992; 8: 115–122.
58. Rutledge DN. Sensitivity and specificity of lump detection in breast models. **Am J Prev Med** 1992; 8: 314–318.
59. Henning KJ, Pollack DM, Freidman SM. A neonatal hepatitis B surveillance and vaccination program: New York City, 1987 to 1988. **Am J Public Health** 1992; 82: 885–888.
60. Rosenbaum P, O'Shea R. Large-scale study of Freedom From Smoking clinics—Factors in quitting. **Public Health Rep** 1992; 107: 150–155.
61. Cohen DA, MacKinnon DP, Dent C, Mason HC, Sullivan E. Group counselling at STD clinics to promote use of condoms. **Public Health Rep** 1992; 107: 727–731.
62. Sacks JJ, Brantley MD, Holmgreen P, Rochat RW. Evaluation of an intervention to reduce playground hazards in Atlanta child-care centers. **Am J Public Health** 1992; 82: 429–431.
63. Graveley EA, Littlefield JH. A cost-effectiveness analysis of three staffing models for the delivery of low-risk prenatal care. **Am J Public Health** 1992; 82: 180–184.
64. Strauss RP, Corless IB, Luckey JW, van der Horst CM, Dennis BH. Cognitive and attitudinal impacts of a university AIDS course: Interdisciplinary education as a public health intervention. **Am J Public Health** 1992; 82: 569–572.
65. Edwards N, Ciliska D, Halbert T, Pond M. Health promotion and health advocacy for and by immigrants enrolled in English as a Second Language classes. **Can J Public Health** 1992; 83: 159–162.
66. Warnecke RB, Langenberg P, Wong SC, Flay BR, Cook TD. The second Chicago televised smoking cessation program: A 24-month follow-up. **Am J Public Health** 1992; 82: 835–840.
67. Rogers J, Grower R, Supino P. Participant evaluation and cost of a community-based health promotion program for elders. **Public Health Rep** 1992; 107: 417–425.
68. Johns MB, Hovell MF, Drastal CA, Lamke C, Patrick K. Promoting prevention services in primary care: A controlled trial. **Am J Prev Med** 1992; 8: 135–140.
69. Rosenberg MJ, Davidson AJ, Chen J-H, Judson FN, Douglas JM. Barrier contraceptives and sexually transmitted diseases in women: A comparison of female-dependent methods and condoms. **Am J Public Health** 1992; 82: 669–674.
70. Calsyn DA, Saxon AJ, Freeman G Jr, Whittaker S. Ineffectiveness of AIDS education and HIV antibody testing in reducing high-risk behaviors among injection drug users. **Am J Public Health** 1992; 82: 573–575.
71. Morse JM, Ewing G, Gamble D, Donahue P. The effect of maternal fluid intake on breast milk supply: A pilot study. **Can J Public Health** 1992; 83: 213–216.
72. Williams EI, Greenwell J, Groom LM. The care of people over 75 years old after discharge from hospital: An evaluation of timetabled visiting by health visitor assistants. **J Public Health Med** 1992; 14: 138–144.
73. Rigotti NA, Bourne D, Rosen A, Locke JA, Schelling TC. Workplace compliance with a no-smoking law: A randomized Community Intervention Trial. **Am J Public Health** 1992; 82: 229–235.
74. Jemmott JB III, Jemmott SW, Fong GT. Reductions in HIV risk-associated sexual behaviors among black male adolescents: Effects of an AIDS prevention intervention. **Am J Public Health** 1992; 82: 372–377.
75. Pirie PL, McBride CM, Hellerstedt W, Jeffery RW, Hatsukami D, Allen S, Lando H. Smoking cessation in women concerned about weight. **Am J Public Health** 1992; 82: 1238–1243.
76. Siska M, Jason J, Murdoch P, Yang WS, Donovan RJ. Recall of AIDs public service announcements and their impact on the ranking of AIDS as a national problem. **Am J Public Health** 1992; 82: 1029–1032.
77. Davies BL, Matte-Lewis L, O'Connor AM, Dulberg CS, Drake ER. Evaluation of the "Time to Quit" self-help smoking cessation program. **Can J Public Health** 1992; 83: 19–23.
78. Lando HA, Hellerstedt WL, Pirie PL, McGovern PG. Brief supportive telephone outreach as a recruitment and intervention strategy for smoking cessation. **Am J Public Health** 1992; 82: 41–46.
79. Lerman C, Ross E, Boyce A, McGovern Gorchov P, McLaughlin R, Rimer B, Engstrom P. The impact of mailing psychoeducational materials to women with abnormal mammograms. **Am J Public Health** 1992; 82: 729–730.
80. Clarke M, Clarke S, Jagger C. Social intervention and the elderly. **Am J Epidemiol** 1992; 136: 1517–1523.
81. Maiman LA, Hildreth NG, Cox C, Greenland P. Improving referral compliance after public cholesterol screening. **Am J Public Health** 1992; 82: 804–809.
82. Abernathy TJ, Bertrand LD. Preventing cigarette smoking among children: Results of a four-year evaluation of the PAL program. **Can J Public Health** 1992; 83: 226–229.
83. Cowan JA, Heckerling PS, Parker JB. Effect of a fact sheet reminder on performance of the periodic health examination: A randomized controlled trial. **Am J Prev Med** 1992; 8: 104–109.
84. McCusker J, Stodart AM, Zapka JG, Morrison CS, Zorn M, Lewis BF. AIDS education for drug abusers: Evaluation of short-term effectiveness. **Am J Public Health** 1992; 82: 533–540.
85. Gilbert JR, Wilson DMC, Singer J, Lindsay EA, Willms DG, Best JA, Taylor DW. A family physician smoking cessation program: An evaluation of the role of follow-up visits. **Am J Prev Med** 1992; 8: 91–95.
86. Cousins JH, Rubovits DS, Dunn JK, Reeves RS, Ramirez AG, Foreyt JP. Family versus individually oriented intervention for weight loss in Mexican American women. **Public Health Rep** 1992; 107: 549–555.
87. Christianson JB, Lurie N, Finch M, Moscovice IS, Hartley D. Use of community-based mental health programs by HMOs: Evidence from a Medicaid demonstration. **Am J Public Health** 1992; 82: 790–796.
88. Beresford SAA, Farmer EMZ, Feingold L, Graves KL, Sumner SK, Baker RM. Evaluation of a self-help dietary intervention in a primary care setting. **Am J Public Health** 1992; 82: 79–84.
89. Hall SM, Turnstall CD, Vila KL, Duffy J. Weight gain prevention and smoking cessation: Cautionary findings. **Am J Public Health** 1992; 82: 799–803.

90. Shattuck AL, White E, Kristal AR. How women's adopted low-fat diets affect their husbands. **Am J Public Health** 1992; 82: 1244–1250.
91. Sorensen G, Morris DM, Hunt MK, Hebert JR, Harris DR, Stoddard A, Ockene JK. Work-site nutrition intervention and employees' dietary habits: The Treatwell program. **Am J Public Health** 1992; 82: 877–880.
92. Anderson DM, Duffy K, Hallet CD, Marcus AC. Cancer prevention counselling on telephone helplines. **Public Health Rep** 1992; 107: 278–283.
93. Rimer BK, Resch N, King E, Ross E, Lerman C, Boyce A, Kessler H, Engstrom PF. Multistrategy health education to increase mammography use among women ages 65 and older. **Public Health Rep** 1992; 107: 369–380.
94. El-Bassel N, Schilling RF. 15-Month followup of women methadone patients taught skills to reduce heterosexual HIV transmission. **Public Health Rep** 1992; 107: 500–504.
95. Landis SE, Schoenbach VJ, Weber DJ, Mittal M, Krishan B, Lewis K, Koch GG. Results of a randomized trial of partner notification in cases of HIV infection in North Carolina. **N Engl J Med** 1992; 326: 101–106.
96. Hannah ME, Hannah WJ, Hellman J, Hewson S, Milner R, Willan A, Canadian Multicenter Post-term Pregnancy Trial Group. Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. **N Engl J Med** 1992; 326: 1587–1592.
97. Johnston CC Jr, Miller JZ, Slemenda CW, Reister TK, Hui MS, Christian JC, Peacock M. Calcium supplementation and increases in bone mineral density in children. **N Engl J Med** 1992; 327: 82–87.
98. Villar J, Farnot U, Barros F, Victoria C, Langer A, Belizan JM. Latin American Network for Perinatal and Reproductive Research. A randomized trial of psychosocial support during high-risk pregnancies. **N Engl J Med** 1992; 327: 1266–1271.
99. Czeizel AE, Dudas I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. **N Engl J Med** 1992; 327: 1832–1835.
100. Chapuy MC, Arlot ME, Duboeuf F, Brun J, Crouzet B, Arnaud S, Delmas PD, Meunier PJ. Vitamin D_3 and calcium to prevent hip fractures in elderly women. **N Engl J Med** 1992; 327: 1637–1642.
101. Tonglet R, Burdoux P, Minga T, Emmans A-M. Efficacy of low oral doses of iodized oil in the control of iodine deficiency in Zaire. **N Engl J Med** 1992; 326: 236–241.
102. Avorn J, Soumeral SB, Everitt DE, Ross-Degnan D, Beers MH, Sherman D, Salem-Schatz SR, Fields D. A randomized trial of a program to reduce the use of psychoactive drugs in nursing homes. **N Engl J Med** 1992; 327: 168–173.
103. Murray GD. The task of a statistical referee. **Br J Surg** 1988; 75: 664–667.
104. Altman DG. Randomisation. Essential for reducing bias. **Br Med J** 1991; 302: 1481–1482.
105. Feinstein AR. Scientific standards in epidemiologic studies of the menace of daily life. **Science** 1988; 242: 1257–1263.
106. Grimes D. Randomized controlled trials: "It ain't necessarily so." Editorial. **Obstet Gynecol** 1991; 78: 703–704.
107. Susser M. Editorial: The trials and tribulations of trials—interventions in communities. **Am J Public Health** 1995; 85: 156–158.
108. Syme LS. Lifestyle intervention in clinic-based trials. **Am J Epidemiol** 1978; 108: 87–91.
109. Lipton JP, Hershaft AM. On the widespread acceptance of dubious medical findings. **J Health Soc Behav** 1985; 5: 328–347.
110. Shulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. **JAMA** 1994; 272: 125–128.
111. Mosteller F. Problems of omission in communications. **Clin Pharmacol Ther** 1979; 25: 761–764.
112. Emerson JD, McPeek B, Mosteller F. Reporting clinical trials in general surgical journals. **Surgery** 1984; 95: 572–579.
113. Shulz KF, Chalmers I, Haynes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. **JAMA** 1995; 273: 408–412.
114. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. **N Engl J Med** 1983; 309: 1358–1361.
115. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Medical. **Stat Med** 1989; 8: 441–454.
116. Miller JN, Colditz GA, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Surgical. **Stat Med** 1989; 8: 455–466.
117. Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. **Lancet** 1990; 335: 149–153.
118. Newell DJ. Intention to treat Analysis. **Int J Epidemiol** 1992; 21: 837–841.
119. Simpson JM, Klar N, Donner A. Accounting for cluster randomization: A review of primary prevention trials, 1990 through 1993. **Am J Public Health** 1995; 85: 1378–1383.
120. Meinert CL. Content of reports on clinical trials: a critical review. **Controlled Clin Trials** 1984; 5: 328–347.
121. Freiman J, Chalmers TC, Smith H, Kuebler R. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. **N Engl J Med** 1978; 299: 690–694.
122. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. **JAMA** 1994; 272: 122–124.
123. Institute for Scientific Information Inc. **Science Citation Index**. JCR Journal Citation Reports: A Bibliometric Analysis of Science Journals in the ISI Database. Philadelphia, 1992.
124. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. **Lancet** 1991; 337: 867–872.
125. Mosteller F, Gilbert JP, McPeek B. Reporting standards and research strategies for controlled trials. **Controlled Clin Trials** 1980; 1: 37–58.
126. Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. **JAMA** 1994; 272: 1926–1931.
127. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. **Controlled Clin Trials** 1995; 16: 62–73.
128. Ad Hoc Working Group for the Critical Appraisal of the Medical Literature. A proposal for more informative abstracts of clinical articles. **Ann Intern Med** 1987; 106: 598–604.
129. Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. **Ann Intern Med** 1990; 13: 69–76.