# Epidemiology and reporting of randomised trials published in PubMed journals

*An-Wen Chan, Douglas G Altman*

*Lancet* 2005; 365: 1159–62

**Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, UK** (Prof D G Altman DSc); **and Department of Medicine, University Health Network Toronto, Canada** (A-W Chan MD)

Correspondence to: Dr An-Wen Chan, Department of Medicine, University of Toronto, Toronto, ON M5G 2C4, Canada **anwen.chan@utoronto.ca**

**Although randomised trials are important for evidence-based medicine, little is known about their overall characteristics. We assessed the epidemiology and reporting of methodological details for all 519 PubMed-indexed randomised trials published in December, 2000 (383 [74%] parallel-group, 116 [22%] crossover). 482 (93%) were published in specialty journals. A median of 80 participants (10th–90th percentile 25–369) were recruited for parallel-group trials. 309 (60%) were blinded. Power calculation, primary outcomes, random sequence generation, allocation concealment, and handling of attrition were each adequately described in less than half of publications. The small sample sizes are worrying, and poor reporting of methodological characteristics will prevent reliable quality assessment of many published trials.**

As far as we are aware, a broad review of randomised trials across study designs, specialties, and journals has not been undertaken since 1980.[1] An updated assessment would help to guide present perceptions and future methodological research. We analysed a representative sample of trial publications in 2000 to provide a cross-sectional view of their characteristics and reporting of methodological details.

We used an extended version of the Cochrane search strategy (phase 1)[2] to identify randomised trials published in December, 2000, and indexed on PubMed by July, 2002. Abstracts were initially screened to exclude obvious non-trials, and complete primary reports in the languages AWC could read (English and French) were reviewed for all remaining studies.

We defined a randomised trial as a prospective study assessing health-care interventions in human participants who were randomly allocated to study groups. Studies of cost-effectiveness and diagnostic test properties were excluded. For pragmatic reasons, trial reports were reviewed twice several months apart by one individual (AWC) rather than by two. Trials were classified by journal type, specialty, and intervention. The trial design, number of randomised groups, number of data collection sites, funding sources, and sample size were also recorded. If information about funding sources and number of study sites was unclear from the trial report, we requested clarification from the trialists.

Finally, the reporting of several important methodological details was assessed, as defined in the panel. Descriptions of power calculations and primary outcomes were recorded. With liberal definitions of adequacy (panel), the reporting of blinding and methods of random sequence generation, allocation concealment, and handling of attrition were also noted. Descriptive summary statistics were calculated and stratified by study design.

519 trials were included in the final cohort (figure). Trial characteristics were stratified by study design (table). 136 (70%) of 193 investigators provided further data about funding sources and study centres on request. We will focus here on parallel-group (383 [74%]) and crossover (116 [22%]) trials because they constituted 96% of our sample.

Primary trial reports were published in 271 different journals, with only 13 journals publishing more than five trials in the single month studied. 482 (93%) trial reports were published in specialty journals and 502 (97%) as full reports. The highest number of reports appeared in *Alimentary Pharmacology and Therapeutics* (n=11), *Anesthesia and Analgesia* (n=11), *Journal of Clinical Pharmacology* (n=10), and *The Lancet* (n=10); all were monthly journals in 2000 apart from *The Lancet*

***Panel:* Definitions used to assess reporting of methodological details in publications of randomised trials**

**Power calculation**
Power calculation stated to have been undertaken

**Specification of primary outcomes**
Primary or main outcomes defined explicitly, or an outcome used in power calculation, or a main outcome described explicitly in primary study objectives

**Blinding**
Study participants, outcome assessors, caregivers, or investigators with no knowledge of the participants' group allocation; or the trial stated to be blinded, or single-blind/double-blind/triple-blind*

**Random sequence generation**
Random method described for allocation of participants to study groups, including computer-generated sequences, random number tables, and coin tosses

**Allocation concealment**
Method described to prevent the individual enrolling trial participants from knowing or predicting the allocation sequence in advance, including central randomisation or envelopes†

**Handling of attrition‡**
Losses to follow-up enumerated for all study groups, and randomised patients with available data were reported as having been analysed in their assigned groups (intention-to-treat)‡

*Recorded as unblinded if explicitly stated as such, or if blinding was clearly not possible (eg, surgery). †Although methods such as envelopes are known to be fallible, we aimed to be conservative and classified them as adequately reported attempts to conceal allocation. ‡Recorded as not reported if one or neither criterion was reported.

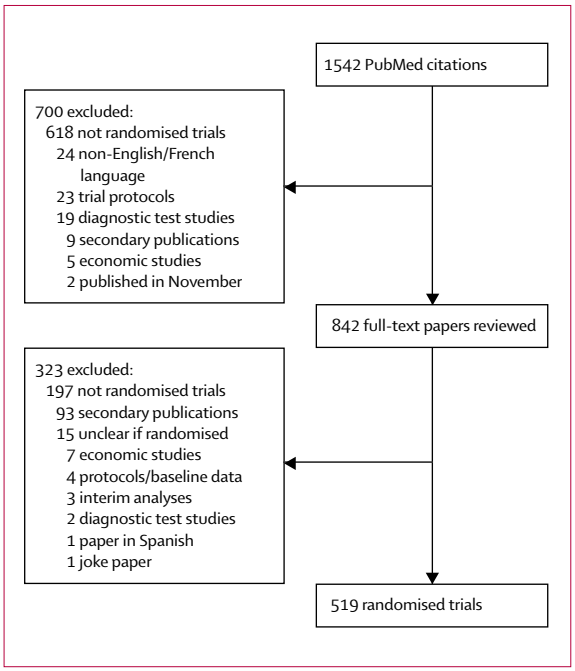


*Figure:* **Identification of randomised trials from PubMed citations of studies published in December, 2000**

(weekly). For parallel-group trials, the most common specialty areas were anaesthetics, paediatrics, cardiology, psychiatry, and gastroenterology (table). Most crossover trials were physiological and pharmacological studies.

More than 70% of parallel-group trials investigated drugs as the primary intervention of interest, whereas more than a tenth assessed procedures or surgical interventions (table). 104 (90%) crossover studies assessed drugs. More than 70% of trials obtained data from one study centre. 97 (25%) parallel-group and 35 (30%) crossover trials compared three or more groups (table).

Parallel-group trials recruited a median of 80 participants overall, with 32 per treatment group (table). Crossover trials randomised a median of 15 participants. With physiological and pharmacokinetic studies excluded, median total sample size was 85 (10th to 90th percentile 26–384) and 20 (10–44) for the remaining 368 parallel-group and 51 crossover trials, respectively.

Sources of funding were reported in publications for 342 (66%) of 519 trials. Investigators for 124 (70%) of the remaining 177 trials provided the following funding details on request: no funding (48 [39%] of 124), non-industry funding (54 [44%] of 124), and part industry funding (22 [18%] of 124). On the basis of publications and survey responses, more than 30% of parallel-group trials received only industry support, with a further 10% reporting part industry support, whereas just over a third had only non-commercial sources of funding (table). Similar proportions were seen for crossover and other study designs.

The table also shows the reporting of methodological details. A power calculation was reported in more than a quarter of publications, whereas primary outcomes were defined in less than half. Blinding was reportedly used in 309 (60%) trials, three-quarters of which were labelled as being double-blinded (229 [74%]). Publications for half the 309 trials used the terms "single-blinded/double-blinded/triple-blinded", "blinded", or "placebo", without providing further details. 90 trials detailed nine different variations of "double-blinding" with respect to who was blinded. Methods of random sequence generation and allocation concealment were each reported in about a fifth of publications. A third of publications adequately detailed the handling of attrition. Quality of reporting was generally worse for crossover studies than for parallel-group trials (table).

This descriptive review provides a cross-sectional epidemiological picture of randomised trials published in December, 2000. Our sample was unrestricted by study design, topic, or journal impact factor, and is representative of the population of PubMed-indexed trials published around the year 2000. However, despite encompassing MEDLINE's coverage of over 4600 journals, the PubMed database might not be representative of all published work. Nonetheless, it is often a primary source of information for researchers and clinicians.

Randomised trials have poor visibility on PubMed. Only a small proportion of over 39 000 citations from December, 2000, were primary reports of randomised trials. Of 1542 citations retrieved in our literature search, two-thirds were ineligible and often needed review of the full paper to ascertain the study design. No journal alone accounted for more than 3% of trials published in the entire month. Only 7% of trials were published in general medical journals. Conclusions from methodological reviews that are limited to trials in these journals might therefore not be representative of published work as a whole.

For comparison with our cohort, two previous reviews used broad MEDLINE samples from 1980[1] and from 1976, 1981, 1986, and 1991.[3] However, comparisons should be interpreted with caution, because these samples differ in characteristics other than the year of publication.

The small sample sizes seen in our cohort are worrying. Little improvement can be seen, relative to the median per treatment group of 19 recorded in 1980,[1] or the 23–39 seen from 1976 to 1991.[3] With the noted median of 32, a two-group comparison has only 39% power to detect a difference between event rates of 10% and 30% at the 0·05 significance level. In practice, this power would be lowered further if losses to follow-up are considered. Trials with inadequate power have a high false-negative error rate and are implicated as a source of publication bias.

| | Trial design | | | |
|---|---|---|---|---|
| | All (n=519) | Parallel-group (n=383) | Crossover (n=116) | Other* (n=20) |
| **Journal type** | | | | |
| Specialty | 482 (93%) | 352 (92%) | 112 (97%) | 18 (90%) |
| General medical† | 37 (7%) | 31 (8%) | 4 (3%) | 2 (10%) |
| **Common specialty fields** | | | | |
| | Physiology 48 (9%) | Anaesthetics 40 (10%) | Physiology 38 (33%) | Dentistry 4 (20%) |
| | Anaesthetics 43 (8%) | Paediatrics 32 (8%) | Pharmacology 27 (23%) | Ophthalmology 4 (20%) |
| | Cardiology 41 (8%) | Cardiology 31 (8%) | Cardiology 10 (9%) | Oncology 3 (15%) |
| | Psychiatry 40 (8%) | Psychiatry 31 (8%) | Endocrinology 7 (6%) | Paediatrics 2 (10%) |
| | Paediatrics 37 (7%) | Gastroenterology 30 (8%) | Psychiatry 7 (6%) | Psychiatry 2 (10%) |
| **Intervention** | | | | |
| Drug | 393 (76%) | 278 (73%) | 104 (90%) | 11 (55%) |
| Surgery/procedure | 51 (10%) | 45 (12%) | 4 (3%) | 2 (10%) |
| Counselling/lifestyle | 55 (11%) | 45 (12%) | 4 (3%) | 6 (30%) |
| Equipment | 20 (4%) | 15 (4%) | 4 (3%) | 1 (5%) |
| **Study centres** | | | | |
| Single | 376 (72%) | 254 (66%) | 106 (91%) | 16 (80%) |
| Multiple | 134 (26%) | 122 (32%) | 9 (8%) | 3 (15%) |
| Unclear | 9 (2%) | 7 (2%) | 1 (1%) | 1 (5%) |
| **Number of study groups** | | | | |
| 2 | 379 (73%) | 286 (75%) | 81 (70%) | 12 (60%) |
| 3 | 85 (16%) | 60 (16%) | 21 (18%) | 4 (20%) |
| 4 | 37 (7%) | 25 (7%) | 9 (8%) | 3 (15%) |
| >4 | 18 (3%) | 12 (3%) | 5 (4%) | 1 (5%) |
| **Sample size** | | | | |
| Median per trial (10th–90th percentile) | 52 (12–310) | 80 (25–369) | 15 (8–38) | 55 (12–556) |
| Median per treatment group per trial (10th–90th percentile) | 32 (12–159) n=393 trials‡ | 32 (12–159) | n/a | 17 (5–202) n=10 trials§ |
| **Funding¶** | | | | |
| Solely industry | 167 (32%) | 122 (32%) | 40 (34%) | 5 (25%) |
| Part industry | 61 (12%) | 39 (10%) | 19 (16%) | 3 (15%) |
| Non-industry | 184 (35%) | 133 (35%) | 41 (35%) | 10 (50%) |
| None | 54 (10%) | 45 (12%) | 8 (7%) | 1 (5%) |
| Unknown | 53 (10%) | 44 (11%) | 8 (7%) | 1 (5%) |
| **Power calculation** | | | | |
| Stated | 142 (27%) | 122 (32%) | 16 (14%) | 4 (20%) |
| Not stated | 377 (73%) | 261 (68%) | 100 (86%) | 16 (80%) |
| **Primary outcome** | | | | |
| Defined | 232 (45%) | 189 (49%) | 36 (31%) | 7 (35%) |
| Not defined | 287 (55%) | 194 (51%) | 80 (69%) | 13 (65%) |
| **Blinding** | | | | |
| Any blinding | 309 (60%) | 214 (56%) | 85 (74%) | 10 (50%) |
| Details provided‖ | 148 (48%) | 114 (53%) | 28 (33%) | 6 (60%) |
| No details provided** | 161 (52%) | 100 (47%) | 57 (67%) | 4 (40%) |
| Unblinded | 166 (32%) | 132 (34%) | 25 (22%) | 9 (45%) |
| Unclear | 44 (8%) | 37 (10%) | 6 (5%) | 1 (5%) |
| **Method of random sequence generation** | | | | |
| Reported | 109 (21%) | 91 (24%) | 11 (9%) | 7 (35%) |
| Computer | 66 (61%) | 57 (63%) | 5 (45%) | 4 (57%) |
| Random number table | 33 (30%) | 28 (31%) | 5 (45%) | 0 |
| Coin toss | 5 (5%) | 2 (2%) | 0 | 3 (43%) |
| Other | 5 (5%) | 4 (4%) | 1 (9%) | 0 |
| Not reported | 410 (79%) | 292 (76%) | 105 (91%) | 13 (65%) |
| **Method of allocation concealment** | | | | |
| Reported | 94 (18%) | 85 (22%) | 7 (6%) | 2 (10%) |
| Envelopes | 48 (51%) | 45 (53%) | 3 (43%) | 0 |
| Central | 25 (27%) | 24 (28%) | 0 | 1 (50%) |
| Pharmacy | 15 (16%) | 11 (13%) | 4 (57%) | 0 |
| Other | 6 (6%) | 5 (6%) | 0 | 1 (50%) |
| Not reported | 425 (82%) | 298 (78%) | 109 (94%) | 18 (90%) |
| **Handling of attrition** | | | | |
| Reported | 174 (34%) | 120 (31%) | 45 (39%) | 9 (45%) |
| Not reported | 345 (66%) | 263 (69%) | 71 (61%) | 11 (55%) |

Data are number (%) of studies unless otherwise indicated. n/a=not applicable. *Includes split-body (n=9), cluster (n=6), factorial (n=4), and n of 1 (n=1) randomised trials. †Journals publishing articles from any specialty. ‡Exclusion of crossover, split-body, and n of 1 trials. §Exclusion of split-body and n of 1 trials. ¶On the basis of information from publications and survey of authors. ‖Reported exactly who was blinded. **Trial was stated to be blinded or single-blinded/double-blinded/triple-blinded (or similar terminology), or placebo was used, without further details provided.

***Table:*** **Characteristics and reporting of methodological details for 519 randomised trials published in December, 2000**

The poor reporting of trial characteristics remains comparable with that seen for trials published in 1980,[1] and prevents reliable assessment of trial methods and bias. The features we examined are included in the revised CONSORT statement for the reporting of parallel-group trials.[4] Information on funding sources was not provided in a third of publications, compared with 44% in 1980.[1] Most trials failed to specify their primary outcomes, which is consistent with the MEDLINE cohort from 1980[1] and is better than the 73% recorded in three general medical journals in 1985.[5] Without explicitly stated primary outcomes that are defined a priori, trialists have free reign to selectively report outcomes, depending on results of significance tests rather than on predefined clinical importance.

The reporting of power calculations was also largely inadequate, with almost three-quarters of trials failing to mention one. This finding represents a mild improvement over the 1980 cohort (96%),[1] and is consistent with trials published in general medical journals in 1979 (88%)[6] and 1994 (68% of negative parallel-group trials).[7] The reporting of an a-priori power calculation can be an indicator of adequate trial planning, help to identify the primary outcome, and indicate whether trials were ended earlier than planned.

The reporting on the presence or absence of blinding (>90%) was much the same as that seen in 1980.[1] However, most publications simply used the ambiguous terms "single-blinded/double-blinded/triple-blinded".[8] Publications should state specifically who was blinded during the trial.[4,8]

Only a fifth of trial publications described the method of random sequence generation, which is comparable with those in 1979 and 1980.[1,6] The frequency of inadequate reporting of allocation concealment in our cohort (82%) is higher than that for trials from general medical journals published in 1994 (63%) and 1998 (46%).[9] Inadequate or unclear reporting of methods of allocation concealment is associated with 30% larger effect estimates.[10]

Our review emphasises design characteristics and deficiencies in the population of published randomised trials. Sample sizes remain low, and much improvement remains to be made in the reporting of methodological details. Introduction of the revised CONSORT statement in 2001[4] and its adoption by journals should help to address the reporting deficiencies identified.

#### Contributors

A-W Chan is the guarantor of the study, had full access to all the data in the study, and is responsible for the integrity of the data and the accuracy of the data analysis. A-W Chan contributed to the study conception and design, acquisition of data, analysis and interpretation of data, and writing of the report. D G Altman contributed to the study conception and design, analysis and interpretation of data, and writing of the report.

#### Conflict of interest statement

We declare that we have no conflict of interest. AWC was funded by the Rhodes Trust; DGA is funded by Cancer Research UK.

#### Acknowledgments

We thank M Clarke and J Deeks for their helpful comments.

#### References

1 Meinert CL, Tonascia S, Higgins K. Content of reports on clinical trials: a critical review. *Control Clin Trials* 1984; **5:** 328–47.
2 Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002; **31:** 150–53.
3 Mulward S, Gøtzsche PC. Sample size of randomized double-blind trials 1976–1991. *Dan Med Bull* 1996; **43:** 96–98.
4 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; **357:** 1191–94.
5 Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987; **317:** 426–32.
6 DerSimonian R, Charette LJ, McPeek B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982; **306**: 1332–37.
7 Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994; **272:** 122–24.
8 Devereaux PJ, Manns BJ, Ghali WA, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001; **285:** 2000–03.
9 Moher D, Jones A, LePage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001; **285:** 1992–95.
10 Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001; **323:** 42–46.