Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study

Lesley Wood, research student,[1] Matthias Egger, head of department and professor of epidemiology and public health,[2] Lise Lotte Gluud, senior registrar,[3] Kenneth F Schulz, vice president of quantitative sciences and clinical professor,[4] Peter Jüni, head of division and reader in clinical epidemiology,[2] Douglas G Altman, director and professor of statistics in medicine,[5] Christian Gluud, head of department,[3] Richard M Martin, reader in clinical epidemiology,[1] Anthony J G Wood, research assistant,[1] Jonathan A C Sterne, professor of medical statistics and epidemiology[1]

[1] Department of Social Medicine, University of Bristol, Bristol BS8 2PR

[2] Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland

[3] Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[4] Family Health International and Department of Obstetrics and Gynaecology, School of Medicine, University of North Carolina at Chapel Hill, NC, USA

[5] Centre for Statistics in Medicine, Oxford, UK

Correspondence to: J A C Sterne
jonathan.sterne@bristol.ac.uk

ABSTRACT

Objective To examine whether the association of inadequate or unclear allocation concealment and lack of blinding with biased estimates of intervention effects varies with the nature of the intervention or outcome.

Design Combined analysis of data from three meta-epidemiological studies based on collections of meta-analyses.

Data sources 146 meta-analyses including 1346 trials examining a wide range of interventions and outcomes.

Main outcome measures Ratios of odds ratios quantifying the degree of bias associated with inadequate or unclear allocation concealment, and lack of blinding, for trials with different types of intervention and outcome. A ratio of odds ratios <1 implies that inadequately concealed or non-blinded trials exaggerate intervention effect estimates.

Results In trials with subjective outcomes effect estimates were exaggerated when there was inadequate or unclear allocation concealment (ratio of odds ratios 0.69 (95% CI 0.59 to 0.82)) or lack of blinding (0.75 (0.61 to 0.93)). In contrast, there was little evidence of bias in trials with objective outcomes: ratios of odds ratios 0.91 (0.80 to 1.03) for inadequate or unclear allocation concealment and 1.01 (0.92 to 1.10) for lack of blinding. There was little evidence for a difference between trials of drug and non-drug interventions. Except for trials with all cause mortality as the outcome, the magnitude of bias varied between meta-analyses.

Conclusions The average bias associated with defects in the conduct of randomised trials varies with the type of outcome. Systematic reviewers should routinely assess the risk of bias in the results of trials, and should report meta-analyses restricted to trials at low risk of bias either as the primary analysis or in conjunction with less restrictive analyses.

INTRODUCTION

Randomised controlled trials, and systematic reviews of these trials, are the best way to evaluate the effects of healthcare interventions.[1] However, defects in the methodological quality of randomised controlled trials may bias their results, and hence the results of meta-analyses that include them.[2] Empirical evidence on specific trial characteristics associated with bias in estimates of intervention effect has come from collections of meta-analyses assembled in so called meta-epidemiological studies.[3]

Several studies have found that trials with inadequate allocation concealment or lack of blinding (see box 1) tend to exaggerate estimates of intervention effects, compared with adequately concealed or adequately blinded trials.[4-7] However, other studies have failed to confirm these associations.[8 9] A possible reason for these inconsistencies is that the magnitude of bias varies between meta-analyses according to other trial characteristics such as the nature of the intervention or outcome. For example, biases associated with lack of blinding might be greater in trials with subjectively assessed outcomes, while bias associated with inadequate allocation concealment may be greater when patients' prognosis can readily be assessed at the time of recruitment to trials, leading to the possibility of selection bias.

We combined data on a large number of controlled trials from three meta-epidemiological studies to examine whether the effects of reported allocation concealment and blinding on estimates of intervention effects differ depending on the nature of the outcome (objective or subjective, all cause mortality or other outcome) and according to whether the meta-analysis examined a drug or non-drug intervention.

METHODS

After a meeting held in November 2001 to which authors of all published meta-epidemiological studies were invited, five datasets were made available for us to use in a combined study. We excluded data from two of the studies: the data of Moher et al were not available in

a suitable format,[5] whereas McAuley et al did not assess the methodological quality of their included trials.[10] We analysed the data from the three remaining studies, which included information on allocation concealment, blinding, and outcome events.[4 6 7]

Schulz et al used 33 meta-analyses from the Pregnancy and Childbirth Group of the Cochrane Collaboration.[4] Each meta-analysis included at least five trials with a combined total of at least 25 outcome events. Kjaergard et al used 14 meta-analyses from 11 systematic reviews, all of which included at least one trial of at least 1000 participants.[6] Egger et al analysed 122 meta-analyses from the *Cochrane Database of Systematic Reviews* that contained at least five randomised trials.[7] Table 1 summarises the methods of each of these studies, including the way that meta-analyses were selected from within systematic reviews and the trial characteristics examined.

We searched Medline, followed by Embase, in order to assign unique identifiers provided by the literature databases to each meta-analysis and each trial. References not indexed in either database were manually assigned a unique identifier. Using the identifier variable, we identified meta-analyses containing at least one overlapping trial. We then removed duplicate meta-analyses until there was no overlap between the remaining meta-analyses except for a small number of trials that contributed to more than one meta-analysis because they had more than one intervention arm or presented more than one outcome.

Box 1: Definitions of trial characteristics studied in this paper

Allocation concealment: procedures that secure strict implementation of the schedule of random assignments by preventing foreknowledge of forthcoming allocations by study participants or by those recruiting them to the trial

- It is always feasible to conceal allocation
- Failure to conceal allocation may lead to biased selection of participants into intervention groups
- Examples of procedures usually considered adequate include sequentially numbered drug containers of identical appearance; central allocation (including web based or pharmacy controlled randomisation); and sequentially numbered, opaque, sealed envelopes
- Examples of procedures usually considered inadequate include using an open random allocation schedule; assignment envelopes without appropriate safeguards (for example, unsealed or non-opaque or not sequentially numbered); and alternation or rotation

Blinding: procedures that prevent study participants, caregivers, or outcome assessors from knowing which intervention was received

- Blinding of participants and caregivers may not be feasible—for example, in a trial of surgery versus radiotherapy for prostate cancer. In such circumstances it may still be possible to blind assessment of outcomes
- Blinding may reduce the risk that knowledge of the intervention received, rather than the intervention itself, affects outcomes or outcome measurements
- Examples of procedures usually considered adequate include provision of indistinguishable placebo tablets and use of a sham surgical procedure in the control group
- An example of blinded outcome assessment is assessment of medical records to ascertain cause of death by an "outcome adjudication committee" unaware of intervention status

---

Assessment of trial quality

In the study by Schulz et al one researcher, who was blinded to the trial outcome, assessed the methodological quality of the included trials using a detailed classification scheme.[4] In the study by Kjaergard et al assessments were done by two observers who were not blinded to study results.[6] Inter-rater reliability of quality assessments was assessed in 30 randomly selected randomised controlled trials and found to be high (intraclass correlation coefficient 0.96). The study of Egger et al was based on quality assessments by authors of the Cochrane reviews, which were generally done in duplicate by two observers.[7]

Table 2 shows the definitions used in the three studies for concealment of allocation and blinding. Definitions of adequate allocation concealment and blinding were similar in all three studies. Allocation concealment was assessed as adequate, unclear, or inadequate in two studies[4 7] and as adequate or inadequate in the other.[6] Blinding was assessed as present (trials described as double blind and using adequate methods such as identical placebo tablets or including blinding of the person assessing outcome) or absent (blinding not performed or not reported, or distinguishable interventions such as tablets and injections were compared). We assessed inter-study reliability of quality assessment using trials included in more than one study and found it was good (median κ statistic 0.67).

Interventions and outcomes

For each meta-analysis in the final dataset, we classified the type of intervention and the type of outcome. Classifications were finalised before we examined associations with trial characteristics. We coded interventions using the classification of Moja et al[11] (drugs; rehabilitation or psychosocial; prevention or screening; surgery or radiotherapy; communication, organisational, or educational; alternative therapeutic; other) and subsequently dichotomised them as drug or non-drug interventions.

We classified outcomes in two ways: firstly, as objectively or subjectively assessed, and, secondly, as all cause mortality or other outcomes. The definition of objective and subjective outcomes was based on the extent to which outcome assessment could be influenced by investigators' judgment. Objectively assessed outcomes included all cause mortality, measures based on a recognised laboratory procedure (such as measurement of haemoglobin concentrations), other objective measures (such as preterm birth), and surgical or instrumental outcomes (all of these were concerned with childbirth, such as caesarean section or instrumental delivery). Note that such surgical outcomes (classified as objectively assessed) depend on doctors' decisions, which could, in the absence of blinding, be affected by knowledge of the intervention received.

Subjectively assessed outcome measures included patient reported outcomes, physician assessed disease outcomes (such as vascular events, pyelonephritis, or respiratory distress syndrome), measures combined

Table 1 | Overview of contributing meta-epidemiological studies

Study	Source of meta-analyses	Choice of meta-analysis within systematic review	Trial characteristics examined	No of meta-analyses (included in current study)	No of trials (included in current study)
Schulz et al 1995[4]	Meta-analyses from Cochrane Pregnancy and Childbirth Group with ≥5 trials containing ≥25 events in the control group, and ≥1 trials with and ≥1 trials without adequate allocation concealment	Meta-analyses with the most homogeneous group of interventions	Allocation sequence generation, allocation concealment, blinding, reporting of exclusions	33 (27)	250 (213)
Kjaergard et al 2001[7]	Meta-analyses in the *Cochrane Library*, Medline or PubMed with at least one trial with ≥1000 patients	Outcome measure described as the primary outcome by the review authors or reported by the largest number of trials	Allocation sequence generation, allocation concealment, blinding, description of dropouts and withdrawal	14 (7)	190 (98)
Egger et al 2003[8]	Meta-analyses from Cochrane Database of Systematic Reviews that had performed comprehensive literature searches	Outcome measure described by the largest number of trials	Publication status, publication in a language other than English, publication in Medline indexed journals, allocation concealment, blinding	122 (112)	1175 (1035)

from several outcomes, and withdrawals or study dropouts. When different methods of outcome assessment were used in different trials in the same meta-analysis we classified the review according to the most subjective method. For example, reviews of smoking cessation used "the most rigorous assessment reported by each included trial." For some trials this was an objective measurement of exhaled carbon monoxide, for some it was repeated questionnaires, and for some it was a single interview. We therefore classified the outcome in this meta-analysis as "patient reported," based on the trials using interviews to assess smoking behaviour.

Statistical methods

We measured intervention effects as odds ratios. Outcome events were recoded where necessary so that an odds ratio below 1 indicated a beneficial effect of the experimental intervention. We calculated the combined effect estimates separately in trials with and without the characteristic of interest (inadequate or unclear allocation concealment or lack of blinding). We used logistic regression models described previously[12] to estimate ratios of odds ratios comparing intervention effects in trials with and without the characteristic of interest. For example, a ratio of odds ratios of 0.7 for trials without blinding would imply that the estimates of intervention effects were exaggerated by 30% in trials without blinding compared with trials with blinding. We derived 95% confidence intervals using robust standard errors allowing for heterogeneity between meta-analyses.[12]

We also calculated ratios of odds ratios separately in each meta-analysis and combined these using random effects meta-analyses, in order to estimate variability between meta-analyses in the effect of trial characteristics.[12] Ratios of odds ratios estimated in this way were consistent with those estimated by logistic regression and are not reported here. Note that meta-analyses in which all trials had the same value of a characteristic (for example, allocation was inadequately concealed in all trials) did not contribute to the estimated effect of that characteristic.

We included interaction terms in logistic regression models to assess whether effects of trial quality varied with the type of intervention or type of outcome. To assess whether there was confounding between the effects of allocation concealment and blinding, we estimated the effect of each characteristic on intervention effects in the same logistic regression model.

Table 2 | Summary of definitions used in meta-epidemiological studies for assessments of allocation concealment of study participants and of blinding

Study	Concealment of allocation			Blinding		
	Adequate	Unclear	Inadequate	Present	Unclear	Absent
Schulz et al[4]	Central randomisation; numbered or coded bottles or containers; drugs prepared by pharmacy; serially numbered, opaque, sealed envelopes; or other description that contained elements convincing of concealing	Not reported or an approach that did not fall into one of the other categories	Alternation or allocation by case record number or date of birth	Described as double blind	NA	Double blinding not reported
Kjaergard et al[7]	Central independent unit, sealed envelopes, or similar	NA	Not described, or open table of random numbers, or similar	Described as double blind, and used an identical placebo or similar method	Described as blinded but method not described	Open (not blind)
Egger et al[8]	Central randomisation; numbered or coded bottles or containers; drugs prepared by pharmacy; serially numbered, opaque, sealed envelopes; or other description that contained elements convincing of concealing	Not mentioned	Reported as alternation, open random number tables, etc	Described as double blind or included blinding of person assessing outcome	Not mentioned	Described as open or similar

NA=Not applicable (category not used in study).

All analyses were done in Stata SE version 9.0 (Stata Corporation, College Station, Texas).

RESULTS

Of 159 meta-analyses contained in the three studies that contributed data, we analysed a total of 146, which included 1346 trials (table 1). Table 3 shows the characteristics of these meta-analyses and trials. The commonest interventions were drug treatments (89 (61%) meta-analyses and 919 (68%) trials) and prevention or screening interventions (22 (15%) meta-analyses, 161 (12%) trials). The commonest types of outcome were physician assessed outcomes (42 (29%) meta-analyses, 323 (24%) trials) and all cause mortality (27 (18%) meta-analyses, 295 (22%) trials). Outcomes were classified as objectively assessed in 78 (53%) meta-analyses and 718 (53%) trials.

Allocation concealment and estimates of intervention effects

We included 102 meta-analyses in our analysis of associations between allocation concealment and estimates of intervention effects (fig 1). Of the 804 trials in these meta-analyses, 272 (34%) had adequate allocation concealment. Overall, intervention effect estimates were exaggerated by 17% in the trials with inadequate or unclear allocation concealment compared with those with adequate allocation concealment (ratio of odds ratios 0.83 (95% CI 0.74 to 0.93)). The size of the bias varied between meta-analyses (heterogeneity variance 0.11, P<0.001).

Bias associated with inadequate or unclear allocation concealment seemed restricted to trials assessing outcomes other than all cause mortality (ratio of odds ratios 0.76 (95% CI 0.66 to 0.87)): there was no evidence of such bias in trials assessing all cause mortality (ratio of odds ratios 1.01 (0.90 to 1.14), interaction P=0.002). The effect of allocation concealment also varied (interaction P=0.009) according to whether trials had subjectively assessed outcomes (ratio of odds ratios 0.69 (0.59 to 0.82)) or objectively assessed outcomes (0.91 (0.80 to 1.03)). The variability between meta-analyses in bias (see fig 1) was lower when the outcome was all cause mortality than for other types of outcome. There was little evidence (interaction P=0.27) that bias associated with inadequate allocation concealment differed in trials that assessed drug interventions compared with trials assessing non-drug interventions.

Blinding and estimates of intervention effects

Figure 2 shows the associations between blinding and estimates of intervention effects, based on 76 meta-analyses containing 746 trials, of which 432 (58%) were blinded. Overall, estimates of intervention effects were exaggerated by 7% in non-blinded compared with blinded trials (ratio of odds ratios 0.93 (0.83 to 1.04)), although the size of the bias varied between meta-analyses (heterogeneity variance 0.11, P<0.001).

The bias associated with lack of blinding was greater (interaction P=0.011) in trials assessing outcomes other than all cause mortality (ratio of odds ratios 0.83 (0.70 to 0.98)) than in those assessing all cause mortality (1.04 (0.95 to 1.14)). Similarly, bias associated with lack of blinding seemed restricted to trials with subjectively assessed outcomes (ratio of odds ratios 0.75 (0.61 to 0.93)), with no evidence of such bias in trials with objective outcomes (1.01 (0.92 to 1.10)) (interaction P=0.010). There was little evidence (interaction P=0.66) that the size of the bias associated with lack of blinding differed between trials of drug and non-drug interventions. The variability between meta-analyses in bias (see fig 2) was lower when the outcome was all cause mortality than for other types of outcome.

It can be argued that, unless allocation is adequately concealed, it is impossible for a trial to be adequately blinded. We therefore restricted analyses of the association of lack of blinding with intervention effect estimates to the trials assessed as having adequate allocation concealment. In this analysis (12 meta-analyses including 60 trials) there was no evidence of an association between blinding and intervention effect estimates (ratio of odds ratios 1.02 (0.92 to 1.14)). Results were similar when further restricted to the seven meta-analyses (42 trials) that assessed objective outcomes (1.03 (0.92 to 1.16)). In the five meta-analyses

Table 3 | Details of interventions and outcomes examined in meta-analyses and trials. Values are numbers (percentages) unless stated otherwise
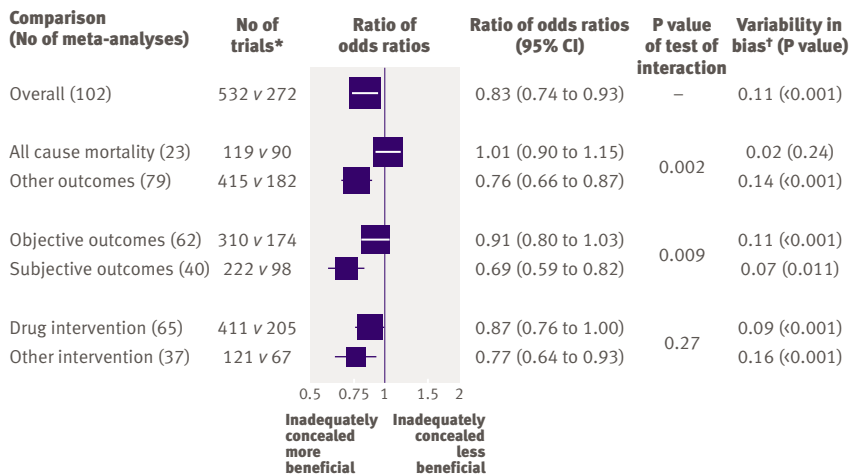
Characteristics	Meta-analyses (n=146)	Clinical trials (n=1346)
Median No of trials (range)	7 (3-67)	—
Median No of participants (range)	1315 (117-44102)	104 (2-17187)
Median year of publication (range)	1998 (1990-2000)	1987 (1962-1998)
Interventions*		
Drug	89 (61)	919 (68)
Rehabilitation or psychosocial	8 (6)	61 (5)
Prevention or screening	22 (15)	161 (12)
Surgery or radiotherapy	5 (3)	36 (3)
Communication, organisational, or educational	2 (1)	28 (2)
Alternative medicine intervention	2 (1)	18 (1)
Other†	18 (12)	123 (9)
Outcomes		
Objective outcomes:		
All cause mortality	27 (18)	295 (22)
Laboratory measurement	17 (12)	144 (11)
Surgical/instrumental‡	16 (11)	159 (12)
Other§	18 (12)	120 (9)
Subjective outcomes:		
Patient reported outcome	16 (11)	207 (15)
Physician assessed disease outcome	42 (29)	323 (24)
Combined outcome	1 (0.7)	11 (1)
Withdrawals or dropouts	9 (6)	87 (6)

*Classification of interventions based on Moja et al.[11]
†Amnioinfusion, amniotomy, holding chambers for β agonist treatment, Doppler ultrasound, electroconvulsive therapy, external cephalic version, enteral nutrition in preterm infants, nasal continuous positive airways pressure in low birthweight infants, timing of induction, suture material, nerve stimulation for low back pain, positioning of umbilical artery catheters, heparinisation of umbilical artery catheters, electronic heart rate monitoring, electronic fetal monitoring plus scalp sampling, birth chair during second stage of labour.
‡Caesarean section (12), instrumental delivery (2), epidural analgesia (1), manual removal of placenta (1).
§Based on birth weight (5), timing of delivery (3), haemorrhage or blood loss measurement (2), non-cephalic birth (2), continuing lactation one week after birth (1), deep venous thrombosis (1), live birth (1), failed delivery (1), episiotomy (1), retention in school grade (1).
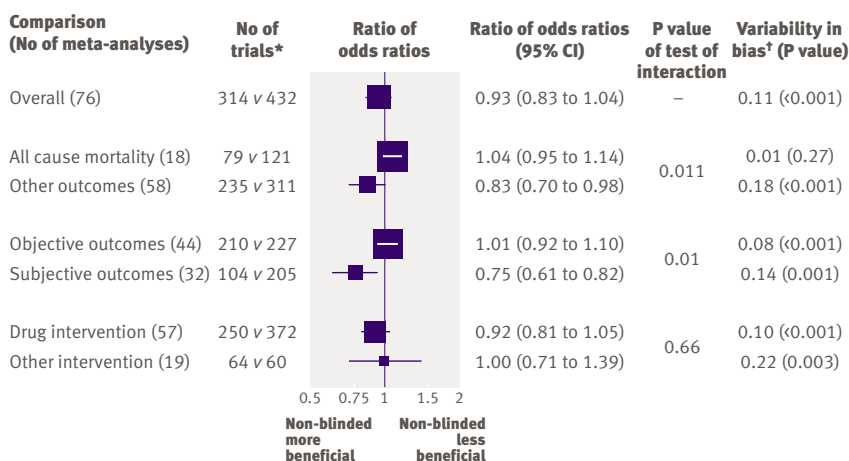
Comparison (No of meta-analyses)	No of trials*	Ratio of odds ratios	Ratio of odds ratios (95% CI)	P value of test of interaction	Variability in bias† (P value)
Overall (102)	532 v 272		0.83 (0.74 to 0.93)	–	0.11 (<0.001)
All cause mortality (23)	119 v 90		1.01 (0.90 to 1.15)	0.002	0.02 (0.24)
Other outcomes (79)	415 v 182		0.76 (0.66 to 0.87)		0.14 (<0.001)
Objective outcomes (62)	310 v 174		0.91 (0.80 to 1.03)	0.009	0.11 (<0.001)
Subjective outcomes (40)	222 v 98		0.69 (0.59 to 0.82)		0.07 (0.011)
Drug intervention (65)	411 v 205		0.87 (0.76 to 1.00)	0.27	0.09 (<0.001)
Other intervention (37)	121 v 67		0.77 (0.64 to 0.93)		0.16 (<0.001)

```
        0.5  0.75  1   1.5  2
        Inadequately    Inadequately
        concealed       concealed
        more            less
        beneficial      beneficial
```

* Inadequately or unclearly concealed *v* adequately concealed
† Between-meta-analysis heterogeneity variance

Fig 1 | Ratios of odds ratios comparing estimates of intervention effects in 532 trials with inadequate or unclear allocation concealment versus 272 trials with adequate concealment. Columns on the right show P values from tests of interaction, and estimates of the variability between meta-analyses in the size of bias

(18 trials) with subjective outcomes, results were similar to those in the main analyses presented in figure 2 (ratio of odds ratios 0.80 (0.49 to 1.31)), although the confidence interval was wide.

Combined analysis of the effects of allocation concealment and blinding

To investigate whether allocation concealment and blinding were independent of each other in their effects on estimates of intervention effect, we included the effects of each characteristic in the same logistic regression model. For all cause mortality and other objectively assessed outcomes, there was little change in the estimated effects of allocation concealment and blinding. For subjectively assessed outcomes, the effect

Comparison (No of meta-analyses)	No of trials*	Ratio of odds ratios	Ratio of odds ratios (95% CI)	P value of test of interaction	Variability in bias† (P value)
Overall (76)	314 v 432		0.93 (0.83 to 1.04)	–	0.11 (<0.001)
All cause mortality (18)	79 v 121		1.04 (0.95 to 1.14)	0.011	0.01 (0.27)
Other outcomes (58)	235 v 311		0.83 (0.70 to 0.98)		0.18 (<0.001)
Objective outcomes (44)	210 v 227		1.01 (0.92 to 1.10)	0.01	0.08 (<0.001)
Subjective outcomes (32)	104 v 205		0.75 (0.61 to 0.82)		0.14 (0.001)
Drug intervention (57)	250 v 372		0.92 (0.81 to 1.05)	0.66	0.10 (<0.001)
Other intervention (19)	64 v 60		1.00 (0.71 to 1.39)		0.22 (0.003)

```
        0.5  0.75  1   1.5  2
        Non-blinded     Non-blinded
        more            less
        beneficial      beneficial
```

* Non-blinded *v* blinded
† Between-meta-analysis heterogeneity variance

Fig 2 | Ratios of odds ratios comparing intervention effect estimates in 314 non-blinded trials versus 432 blinded trials. Columns on the right show P values from tests of interaction, and estimates of the variability between meta-analyses in the size of bias

of each characteristic was modestly attenuated (ratio of odds ratios for inadequate or unclear allocation concealment 0.75 (0.63 to 0.88), for lack of blinding 0.77 (0.65 to 0.91)). Similar attenuation was seen for outcomes other than all cause mortality (ratio of odds ratios for inadequate or unclear allocation concealment 0.80 (0.73 to 0.87), for lack of blinding 0.85 (0.76 to 0.95)).

DISCUSSION

We analysed a dataset that included many meta-analyses and trials and represented a wide variety of clinical disciplines, interventions, and outcomes. We found that bias in intervention effect estimates from clinical trials, resulting from inadequate allocation concealment and lack of blinding, varied according to the type of outcome assessed. There was little evidence of bias in trials assessing all cause mortality or other objectively assessed outcomes. In contrast, inadequate allocation concealment and lack of blinding were associated with over-optimistic estimates of intervention effects for subjectively assessed outcomes. The size of bias varied between meta-analyses, but this variability was less in trials assessing all cause mortality, where there was little evidence of bias. It thus seems that in situations when bias is likely the size of such bias is unpredictable.

Strengths and weaknesses of study

The reported quality of a trial may not reflect the way that it was conducted in practice: well conducted trials may be reported badly.[13 14] For example, interpretations and definitions of single, double, and triple blinding have been shown to vary greatly between physicians and textbooks.[15] In our study many trials will have been classified as blinded solely because they were described as "double blind" without any further details on who was blinded and how blinding was achieved. Even clearer patterns than those identified here might have emerged had more accurate information on the conduct of the trials been available.

We assembled a large dataset by combining data from three meta-epidemiological studies, but the way in which trial characteristics were assessed varied between the original studies. However, cross classification of trials that were assessed in at least two studies indicated good agreement. The study of Egger et al[7] relied on assessments of trial quality by the authors of Cochrane reviews rather than by methodological experts. Despite the standardised guidelines specified in the Cochrane handbook,[16] evaluations by authors of Cochrane reviews of whether a study had adequate allocation concealment may be inconsistent.[17] The effect of trial quality on estimates of intervention effect in this study was, however, in line with previous studies in which quality was assessed by the same observers: one would expect attenuation of effects if assessments in Cochrane reviews were less reliable.

Other types of bias that may affect the results of randomised controlled trials include bias due to selective reporting of outcomes[18 19] and bias due to

exclusion of patients from analyses because of lack of adherence to the randomised intervention or missing outcome data.[20] We did not account for these sources of bias in our study, and hence our assessment of the effects of inadequate allocation concealment or lack of blinding could be confounded by them, as well as by characteristics of study populations and interventions that influence estimates of intervention effects.

Results in context with other literature

Some meta-epidemiological studies, including but not restricted to those combined here, have provided evidence that inadequate or unclear allocation concealment and lack of double blinding are associated with over-optimistic estimates of the effect of interventions assessed in clinical trials.[4-7] However, the evidence is not consistent. Balk et al found no association between adequacy of allocation concealment or double blinding and effect estimates.[8] Als-Nielsen et al also found no association with adequacy of allocation concealment, but they did find that trials with inadequate or unclear generation of the allocation sequence tended to exaggerate effects.[9 21] Their estimated effect of blinding was similar to that reported in our study.

Interestingly, the largest group of meta-analyses included in the study of Balk et al[8] were of interventions against cardiovascular disease, specifically selected because the outcome was all cause mortality. This may explain why they found little evidence of an effect of either allocation concealment or blinding. For the other topics examined by Balk et al (infectious disease, paediatrics, and surgery), confidence intervals for the effects of allocation concealment and blinding were

---

WHAT IS ALREADY KNOWN ON THIS TOPIC

Flaws in the conduct of randomised controlled trials, such as inadequate allocation concealment or lack of blinding, may bias their results

Several studies have found that such flaws lead to exaggerated estimates of the effects of medical interventions, but other studies have failed to confirm these associations

WHAT THIS STUDY ADDS

The average bias associated with lack of adequate allocation concealment or lack of blinding was less for trials with objectively assessed outcomes than trials with subjectively assessed outcomes, and less for trials with all cause mortality as the outcome than for trials with other outcomes

Except when the outcome was all cause mortality, the size of bias varied between meta-analyses. Therefore, in situations when bias is likely the size of such bias is unpredictable

Since the bias associated with inadequate or unclear allocation concealment varied with outcome, it may result in part from an association with subsequent flaws in the conduct of the trial rather than from selection biases. The effect of allocation concealment remained even after adjustment for blinding, suggesting it is a marker for other bias-reducing strategies, beyond blinding

---

wide and consistent with the overall effects reported here, although analyses stratified by type of outcome were not presented.

Possible mechanisms

Our finding that bias resulting from inadequate allocation concealment varied with the type of outcome variable was unexpected. The purpose of allocation concealment is to avoid selection bias, whereby knowledge of prognosis at the time of recruitment to a randomised trial influences the intervention group to which the patient is allocated.[17 22] Such selection bias would be expected to be greatest when it is easy to assess patients' prognosis at the time they are recruited to a trial, and affect the results of trials with objectively assessed as well as subjectively assessed outcomes. Conversely, lack of blinding would be expected to introduce bias if knowledge of intervention groups affected the care received or the assessment of outcomes. This could happen independently of whether selection bias due to inadequate concealment of allocation was present.

There was a strong association between inadequate or unclear allocation concealment and lack of blinding: about three quarters of trials with adequate concealment were classified as blinded, compared with half of trials with inadequate or unclear concealment. In trials with inadequate or unclear allocation concealment or lack of double blinding, exaggerated intervention effects continued to be seen for subjectively assessed outcomes when the effect of each characteristic was estimated after controlling for the other. Only 60 trials, in 12 meta-analyses, contributed to analyses of the effect of blinding when these were restricted to trials with adequate concealment of allocation. We were thus unable to estimate with precision the effect of blinding independently of the effect of allocation concealment. The bias we observed with inadequate or unclear allocation concealment may result at least in part from its association with subsequent flaws in the conduct of the trial, in particular with biased outcome assessment, rather than from selection bias. Adequate allocation concealment may also be a marker for other strategies for reducing bias, beyond blinding, as its effect remains when adjusted for blinding.

Implications

There is a broad consensus that the reporting of randomised controlled trials should be improved and standardised, with many medical journals having adopted the CONSORT (consolidated standards of reporting trials) recommendations.[23 24] Many of the trials included in this study were published before the CONSORT statement. Improved transparency about the way that trials were conducted should allow a refined understanding of the potential of different aspects of the conduct of trials to bias results, and how such bias varies between different settings. Our results suggest that efforts to minimise selection and performance bias are particularly important when objective measurement of outcomes is not feasible. It is always

Box 2: Recommendations for practice

- Since bias associated with lack of adequate allocation concealment or lack of blinding is less for trials with objectively assessed outcomes than trials with subjectively assessed outcomes, efforts to minimise bias are particularly important when objective measurement of outcomes is not feasible
- Authors of systematic reviews, and those critically appraising trials, should routinely assess the risk of bias in results associated with the way each trial was done. Such assessments should be outcome specific. The Cochrane Collaboration has recently formulated detailed guidance on how to do this (see www.cochrane.org/resources/handbook)
- Systematic reviewers should present meta-analyses restricted to trials at low risk of bias for each outcome, either as the primary analysis or in conjunction with less restrictive analyses

possible to conceal the allocation of patients to intervention groups, for example by using a telephone randomisation service. In contrast, blinding of participants and physicians is not always feasible; such as when surgery is compared with radiotherapy for treatment of cancer, or when a drug has unmistakable adverse effects. Our results suggest that in such situations it is desirable to specify at least one objectively assessed outcome, even if the outcome of most interest is subjective.

Systematic reviewers should routinely assess the risk of bias in results associated with the way each trial was done. Such critical appraisal should allow for our finding that lack of blinding or of adequate allocation concealment is particularly important for subjectively assessed outcomes. Some critical appraisal checklists, such as the GATE appraisal,[25] ask whether the outcome measurements are either blinded or objective.

The Cochrane Collaboration is publishing detailed guidance on assessing risk of bias, based on both empirical evidence and expert opinion, available at www.cochrane.org/resources/handbook/.[26] This identifies six key domains of trial conduct and analysis—sequence generation; allocation concealment; blinding of participants, personnel, and outcome assessors; incomplete outcome data; selective outcome reporting; and other factors (specific for study design and topic). This guidance emphasises that assessments of risk of bias due to lack of blinding or incomplete data must be outcome specific. Domain specific assessments need to be combined into an overall assessment of the risk of bias (low, unclear, or high) for each outcome.

There is currently no simple or widely accepted method to integrate assessments of risk of bias into meta-analyses. The variability in bias between meta-analyses that we identified implies that more emphasis in meta-analyses should be given to trials at low risk of bias, and it may be possible to develop statistical methods to do this.[27] Currently, we believe that systematic reviewers should present meta-analyses restricted to trials at low risk of bias for each outcome, either as the primary analysis or in conjunction with less restrictive analyses.

Conclusions

Box 2 summarises our recommendations for practice. The effects on bias in intervention effect estimates of two key aspects of the conduct of randomised controlled trials—adequacy of reported allocation concealment and reported blinding—vary with the type of outcome variable. In trials with subjectively assessed outcomes lack of adequate allocation concealment or of blinding tend to produce over-optimistic estimates of the effect of interventions. Except for trials with all cause mortality as the outcome, the size of bias varies between meta-analyses. Systematic reviewers should routinely assess the risk of bias in the results of included trials (and such assessments should be outcome specific), and should report meta-analyses restricted to trials at low risk of bias either as the primary analysis or in conjunction with less restrictive analyses.

1 Chalmers I, Altman DG. *Systematic reviews*. London: BMJ Publishing Group, 1995.
2 Gluud LL. Bias in clinical intervention research. *Am J Epidemiol* 2006;163:493-501.
3 Naylor CD. Meta-analysis and the meta-epidemiology of clinical research. *BMJ* 1997;315:617-9.
4 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
5 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.
6 Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001;135:982-9.
7 Egger M, Jüni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 2003;7:1-76.
8 Balk EM, Bonis PA, Moskowitz H, Schmid CH, Ioannidis JP, Wang C, et al. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *JAMA* 2002;287:2973-82.
9 Als-Nielsen B, Chen W, Gluud LL, Siersma V, Hilden J, Gluud C. Are trial size and reported methodological quality associated with treatment effects? Observational study of 523 randomised trials. 12th International Cochrane Colloquium, Ottawa, Canada. 2004.

10 McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion on grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* 2000;356:1228-31.

11 Moja LP, Telaro E, D'Amico R, Moschetti I, Coe L, Liberati A. Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study. *BMJ* 2005;330:1053.

12 Sterne JAC, Jüni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med* 2002;21:1513-24.

13 Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA* 2002;287:2801-4.

14 Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swann S, et al. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;328:22-4.

15 Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000-3.

16 Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions. Version 5.0.0. Cochrane Collaboration,* 2008. www.cochrane.org/resources/handbook.

17 Pildal J, Chan AW, Hróbjartsson A, Forfang E, Altman DG, Gotzsche PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. *BMJ* 2005;330:1049-52.

18 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.

19 Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;171:735-40.

20 Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005;34:79-87.

21 Siersma V, Als-Nielsen B, Chen W, Hilden J, Gluud LL, Gluud C. Multivariable modelling for meta-epidemiological assessment of the association between trial quality and treatment effects estimated in randomized clinical trials. *Stat Med* 2007;26:2745-58.

22 Forder PM, Gebski VJ, Keech AC. Allocation concealment and blinding: when ignorance is bliss. *Med J Aust* 2005;182:87-9.

23 Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.

24 Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.

25 Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, et al. The GATE frame: critical appraisal with pictures. *Evid Based Med* 2006;11:35-8.

26 Higgins JPT, Altman DG, eds. Assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions Version 5.0.0. Cochrane Collaboration,* 2008. www.cochrane.org/resources/handbook.

27 Welton NJ, Ades AE, Carlin JB, Altman DG, Sterne JAC. Models for potentially biased evidence in meta-analysis using empirically based priors. *J R Statist Soc A* (in press).