RESEARCH

The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study

Eveline Nüesch, research fellow,[1,2] Sven Trelle, associate director,[1,2] Stephan Reichenbach, senior research fellow,[1,3] Anne W S Rutjes, senior research fellow,[1,4] Elizabeth Bürgi, research fellow,[5] Martin Scherer, professor of health services research,[6,7] Douglas G Altman, professor of statistics in medicine,[8] Peter Jüni, head of division[1,2]

[1]Institute of Social and Preventive Medicine, University of Bern, Switzerland

[2]CTU Bern, Bern University Hospital, Switzerland

[3]Department of Rheumatology, Immunology and Allergology, Bern University Hospital, Switzerland

[4]Department of Clinical Pharmacology and Epidemiology, Consorzio Mario Negri Sud, Santa Maria Imbaro, Chieti, Italy

[5]Department of Internal Medicine, Bern University Hospital, Switzerland

[6]Department of General Practice, University of Göttingen, Germany

[7]Institute of Social Medicine, University of Luebeck, Germany

[8]Centre for Statistics in Medicine, University of Oxford, Oxford

Correspondence to: P Jüni
juni@ispm.unibe.ch

ABSTRACT

Objective To examine whether excluding patients from the analysis of randomised trials are associated with biased estimates of treatment effects and higher heterogeneity between trials.

Design Meta-epidemiological study based on a collection of meta-analyses of randomised trials.

Data sources 14 meta-analyses including 167 trials that compared therapeutic interventions with placebo or non-intervention control in patients with osteoarthritis of the hip or knee and used patient reported pain as an outcome.

Methods Effect sizes were calculated from differences in means of pain intensity between groups at the end of follow-up, divided by the pooled standard deviation. Trials were combined by using random effects meta-analysis. Estimates of treatment effects were compared between trials with and trials without exclusions from the analysis, and the impact of restricting meta-analyses to trials without exclusions was assessed.

Results 39 trials (23%) had included all patients in the analysis. In 128 trials (77%) some patients were excluded from the analysis. Effect sizes from trials with exclusions tended to be more beneficial than those from trials without exclusions (difference −0.13, 95% confidence interval −0.29 to 0.04). However, estimates of bias between individual meta-analyses varied considerably (τ^2=0.07). Tests of interaction between exclusions from the analysis and estimates of treatment effects were positive in five meta-analyses. Stratified analyses indicated that differences in effect sizes between trials with and trials without exclusions were more pronounced in meta-analyses with high between trial heterogeneity, in meta-analyses with large estimated treatment benefits, and in meta-analyses of complementary medicine. Restriction of meta-analyses to trials without exclusions resulted in smaller estimated treatment benefits, larger P values, and considerable decreases in between trial heterogeneity.

Conclusion Excluding patients from the analysis in randomised trials often results in biased estimates of treatment effects, but the extent and direction of bias is unpredictable. Results from intention to treat analyses

should always be described in reports of randomised trials. In systematic reviews, the influence of exclusions from the analysis on estimated treatment effects should routinely be assessed.

INTRODUCTION

In clinical trials, deviations from protocol and losses to follow-up often lead to the exclusion of some randomised patients from the analysis.[1 2] Patients excluded after randomisation are unlikely to be representative of patients remaining in the trial. For example, patients may not be available for follow-up because they have an acute exacerbation of their condition or severe side effects,[3] and patients with protocol deviations may have a worse prognosis than those adhering to the protocol.[4] The selective occurrence and biased handling of protocol deviations and losses to follow-up may lead to results that differ systematically from the true values. This is generally referred to as attrition bias.[2] To ensure that intervention and control groups are comparable and to prevent attrition bias, all randomised patients should be included in the analysis and kept in their original groups, regardless of their adherence to the study protocol. In other words, the analysis should be done according to the intention to treat principle, avoiding any selective exclusion of patients after randomisation.[2 5]

Meta-epidemiological studies examine the association of specific characteristics of a trial, such as concealment of allocation or blinding of patients, with estimated treatment effects in a collection of meta-analyses and their component trials.[6 7] The association of withdrawals, dropouts, and exclusions after randomisation with estimated treatment effects has been explored in four meta-epidemiological studies of binary outcomes.[1 6 8-10] The direction and magnitude of attrition bias varied between different studies according to different methods and definitions used and different clinical areas addressed: attrition bias may result in both overestimation and underestimation of treatment effects, and its magnitude is difficult to predict.[1 9 11 12] In general, randomised controlled trials using subjective outcomes are more susceptible to bias

than trials using objective outcomes such as overall mortality. A recent study found that bias associated with the lack of allocation concealment and lack of blinding was restricted to trials using subjectively assessed outcomes.[12] In trials of osteoarthritis, treatment effects are often evaluated using subjective outcomes, such as intensity of pain or disability measured on visual analogue, numerical rating, or Likert scales, whereas objective binary outcomes, such as mortality, are addressed rarely. Meta-analyses of osteoarthritis trials may therefore be particularly prone to attrition bias associated with exclusions of patients from the analysis.[2]

We carried out a meta-epidemiological study to assess the impact of attrition bias in meta-analyses of non-binary patient reported outcomes, such as pain intensity. We examined whether excluding patients from the analysis were associated with biased estimates of treatment effects and with increased heterogeneity between trials in meta-analyses of interventions used for the treatment of pain in osteoarthritis.

METHODS

Search and selection of meta-analyses and component trials

We searched the Cochrane Library, Medline, Embase, and CINAHL using a combination of keywords and text words related to osteoarthritis. These were combined with validated filters for systematic reviews and meta-analyses.[13] The last update was carried out on 20 November 2007 (see web extra appendix table 1 for details of search strategy).

We included meta-analyses of randomised or quasi-randomised trials in patients with osteoarthritis of the knee or hip. Trials using an unpredictable allocation sequence were considered as randomised, trials using potentially predictable allocation mechanisms, such as alternation or the allocation of patients according to their date of birth, were considered as quasi-randomised. Meta-analyses were eligible if they assessed patient reported pain comparing any intervention with placebo, sham, or a non-intervention control. Two reviewers independently evaluated the reports for eligibility, and disagreements were resolved by discussion. If necessary a third reviewer was consulted to reach consensus. Reports of all component trials were obtained, and no language restrictions were applied.

Data collection and quality assessment

Two reviewers used a standardised form to independently extract data from the original reports of individual trials on interventions, funding, year of publication, publication language, design, study size, blinding of patients, losses to follow-up, exclusions, handling of missing data, and results. When necessary we approximated means and measures of dispersion from figures. For crossover trials we extracted data from the first period only.[14] Disagreements were resolved by discussion with a third reviewer and subsequent consensus.

Trials were classified to have had no exclusions of patients from the analysis if there was an explicit statement that all randomised patients were included in the analysis of the outcome we extracted or if the reported numbers of patients randomised and analysed on this outcome were identical. We classified trials to have had exclusions if they explicitly reported exclusions from the analysis, if the number of patients analysed was lower than the number of patients randomised, or if it was unclear whether exclusions from the analysis had occurred. Concealment of treatment allocation was considered adequate if the investigators responsible for patient inclusion were unable to suspect before allocation which treatment was next—central randomisation or the use of sequentially numbered, sealed, and opaque randomisation envelopes was deemed adequate, for example. Blinding of patients was considered adequate if experimental and control interventions were described as indistinguishable or if a double dummy technique was used.[2]

Outcome measures

The primary outcome was patient reported pain. If different pain outcomes were reported we extracted one outcome per study according to a hierarchy described previously.[15 16] If more then one time point was reported we extracted the latest time point up to three months after the end of treatment for potentially structure modifying agents and up to 12 months after the end of treatment for behaviour changing interventions. For all other interventions we extracted the outcome at the end of treatment.

Statistical analysis

We expressed treatment effects as effect sizes by dividing the difference in mean values at the end of follow-up by the pooled standard deviation. Negative effect sizes indicate a beneficial effect of the experimental intervention. If some required data were unavailable we used approximations as previously described.[16] If a trial was included in more than one meta-analysis we inflated standard errors to avoid double counting of patients—for example, if the control group of a trial with three arms was included in two different meta-analyses, we inflated the standard error of the estimate for the control group by $\sqrt{2}$. We used standard inverse variance random effects meta-analyses to combine effect sizes across trials and calculated the variance estimate τ^2 as a measure of heterogeneity.[17]

Within each meta-analysis we used random effects meta-analysis to estimate effect sizes separately for trials with and trials without exclusions of patients from the analysis. Then we derived differences between estimates from trials with exclusions and trials without exclusions for each meta-analysis and combined these differences using random effects meta-analysis, which fully accounted for the variability in bias between meta-analyses.[7] A negative difference in effect sizes indicates that trials with exclusions show a more beneficial treatment effect. Meta-analyses including only trials with exclusions or only trials without
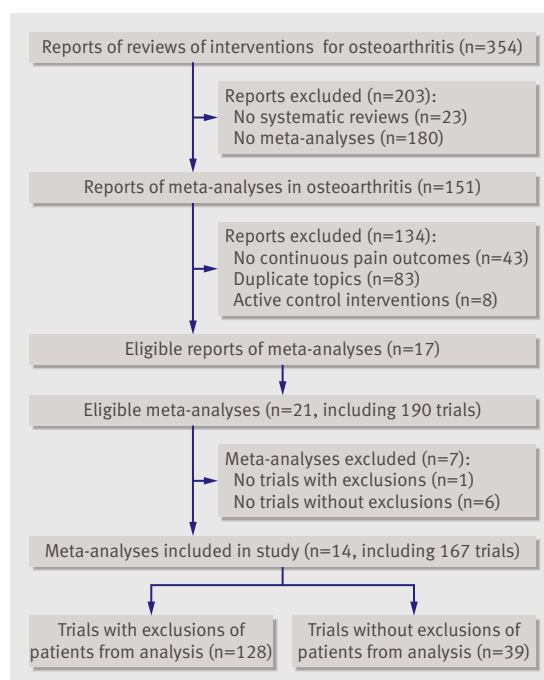
Fig 1 | Identification of meta-analyses in osteoarthritis

exclusions did not contribute to the analysis. Formal tests of interaction between exclusions from the analysis and estimated treatment benefits were done separately for each meta-analysis based on z scores for estimated differences in effect sizes between trials with and trials without exclusions and the corresponding standard errors. We carried out stratified analyses accompanied by interaction tests according to the following characteristics: between trial heterogeneity in the overall meta-analysis (low, $\tau^2<0.06$, v high, $\tau^2\geq0.06$), treatment benefit in the overall meta-analysis (small, effect sizes >-0.5, v large, effect sizes ≤-0.5),[15 18]

and type of intervention assessed in the meta-analysis (drug v other interventions, conventional v complementary medicine). A τ^2 of 0.06 corresponds to a difference between smallest and largest effect sizes of about 1 standard deviation unit.[19] To control confounding by concealment of allocation and by patient blinding, we used stratification by these factors to derive differences between trials with and trials without exclusions adjusted for concealment of allocation and adjusted for patient blinding.

Finally, we compared pooled effect sizes, between trial heterogeneity, precision defined as the inverse of the standard error, and P values for pooled effect sizes between overall random effects meta-analyses including all trials and restricted meta-analyses including trials without exclusions only. Measures were compared using scatter plots and Wilcoxon rank tests for paired observations. P values were two sided. All analyses were done in STATA version 10.

RESULTS

Characteristics of included meta-analyses

Overall, 354 reports of reviews of interventions in osteoarthritis were identified (fig 1). Seventeen reports including 21 meta-analyses were eligible. Of these, 14 meta-analyses included at least one trial with and one without exclusion of patients from the analysis and contributed to the study.[16 20-30] Table 1 shows the characteristics of the included meta-analyses. The meta-analyses included 167 trials in 41 170 patients. Eight meta-analyses assessed the efficacy of drug interventions and five assessed interventions in complementary medicine. The number of trials per meta-analysis ranged from three to 24 (median 11) and the number of patients contributing to the meta-analysis from 278 to 13 659 (median 1731). The pooled effect sizes derived from random effects meta-analyses

Table 1 | Characteristics of included meta-analyses

Interventions	Drug intervention	Complementary medicine	No of trials*	No of patients	Effect size (95% CI)	Heterogeneity τ^2 (P value)
Exercise v control[20]	No	No	16	2700	−0.29 (−0.38 to −0.21)	0.00 (0.41)
Viscosupplementation v placebo[21]	Yes	No	22	3046	−0.33 (−0.50 to −0.17)	0.11 (‹0.001)†
Self management v control[22]	No	No	12	5812	−0.07 (−0.15 to 0.02)	0.01 (0.15)
Glucosamine v placebo[23]	Yes	Yes	15	1518	−0.61 (−0.94 to −0.28)‡	0.35 (‹0.001)†
Diacerein v placebo[25]	Yes	No	6	1613	−0.24 (−0.35 to −0.13)	0.00 (0.33)
Acetaminophen (paracetamol) v placebo[24]	Yes	No	5	1849	−0.23 (−0.37 to −0.10)	0.01 (0.13)
Opioids v placebo[26]	Yes	No	13	3713	−0.39 (−0.47 to −0.31)	0.00 (0.26)
Oral NSAIDs v placebo[29]	Yes	No	24	13659	−0.40 (−0.49 to −0.31)	0.04 (‹0.001)
Topical NSAIDs v placebo[29]	Yes	No	9	1302	−0.47 (−0.65 to −0.29)	0.04 (0.018)
Low-level laser therapy v placebo[28]	No	Yes	8	347	−0.47 (−0.98 to 0.04)	0.42 (‹0.001)†
TENS v sham[28]	No	Yes	10	358	−0.88 (−1.36 to −0.39)‡	0.52 (‹0.001)†
Weight reduction v control[27]	No	No	3	278	−0.12 (−0.33 to 0.09)	0.01 (0.34)
Acupuncture v control[30]	No	Yes	6	1540	−0.49 (−0.78 to −0.19)	0.12 (‹0.001)†
Chondroitin v placebo[16]	Yes	Yes	20	3833	−0.72 (−0.95 to −0.49)‡	0.23 (‹0.001)†

NSAIDs=non-steroidal anti-inflammatory drugs; TENS=transcutaneous electrical nerve stimulation. Effect sizes and corresponding 95% confidence intervals were derived from random effects meta-analyses of all trials. Negative effect sizes indicate a beneficial effect of experimental intervention. Meta-analyses are ordered according to year of publication.
*Number of trials totals 169 as two trials were included each in two different meta-analyses.
†Meta-analyses considered to have high heterogeneity between trials ($\tau^2\geq0.06$).
‡Meta-analyses considered to have large estimated treatment benefit according to overall meta-analysis including all trials (pooled effect size ≤-0.50).

Table 2 | Characteristics of component trials

Characteristics	No (%) of trials with exclusions (n=128)	No (%) of trials without exclusions (n=39)	P value
Losses to follow-up:			
None*	1 (1)	9 (23)	
<10%	32 (26)	6 (15)	
10-20%	26 (20)	6 (15)	0.002
>20%	26 (20)	17 (44)	
Information unavailable	43 (33)	1 (3)	
Imputation of missing data:			
Last observation carried forward	35 (27)	19 (49)	
Multiple imputation	5 (4)	6 (15)	<0.001
Explicitly no losses to follow-up*	1 (1)	8 (21)	
Information unavailable	87 (68)	6 (15)	
Adequate concealment of allocation:			
Yes	24 (19)	15 (38)	0.07
No or unclear	104 (81)	24 (62)	
Described as double blind:			
Yes	88 (69)	28 (72)	0.74
No	40 (31)	11 (28)	
Adequate blinding of patients:			
Yes	60 (46)	21 (54)	0.43
No or unclear	68 (54)	18 (46)	
Primary outcome:			
Reported	70 (55)	27 (69)	0.27
Not reported	58 (45)	12 (31)	
Sample size calculation:			
Reported	50 (39)	23 (59)	0.12
Not reported	78 (61)	16 (41)	
No of patients randomly assigned:			
>200	48 (38)	21 (54)	0.20
≤200	80 (62)	18 (46)	
Drug intervention:			
Yes	85 (66)	28 (72)	0.53
No	43 (34)	11 (28)	
Complementary medicine:			
Yes	44 (34)	15 (38)	0.62
No	84 (66)	24 (62)	
Funding by non-profit organisation:			
Yes	30 (24)	9 (23)	0.96
No or unclear	98 (76)	30 (77)	
Language of primary report:			
English	120 (94)	38 (97)	0.49
Non-English	8 (6)	1 (3)	
Year of publication:			
1980-9	17 (13)	0 (0)	
1990-9	49 (38)	10 (26)	0.003
2000-7	62 (49)	29 (74)	

P values are derived from logistic regression models adjusted for clustering of trials within meta-analyses. Comparisons of frequency of concealment of allocation, description of double blinding, adequate blinding of patients, trial size, type of intervention, funding, language of publication, and publication year were preplanned. *One trial reporting that no patient was lost to follow-up used the last observation carried forward approach to impute some missing outcome data.

including all trials ranged from −0.07, indicating essentially no benefit, to −0.88, representing a large benefit (median −0.40). All meta-analyses favoured the experimental intervention and 11 of 14 showed statistically significant differences between experimental and control intervention at the conventional level of P=0.05. The variance τ^2 as a measure of between trial heterogeneity varied between 0.00 and 0.52 (median 0.04, table 1).

Characteristics of component trials

Table 2 shows the characteristics of included trials. In total, 39 of the 167 trials (23%) included all randomised patients in the analysis. In 114 trials (69%) there were exclusions, and in 14 trials (8%) it was unclear whether exclusions had occurred. Exclusions ranged from 0.1% to 40% (median 7.2%). Trials with exclusions were less likely to provide information on losses to follow-up (P=0.002). Data imputations using the last observation carried forward method were reported by 27% of trials with exclusions and 49% of trials without exclusions, multiple imputation by 4% and 15%, and for 68% and 15% it was unclear how the trialists dealt with missing data in the analysis. Trials with exclusions were published earlier (mean 1998, SD 6) than trials without exclusions (2001, SD 4; P=0.002) and tended to report adequate concealment of allocation and sample size calculations less often. No clear differences were observed for blinding, reporting of primary outcome, type of intervention, source of funding, and language of publication.

Effect of exclusions on estimates of treatment effects

Figure 2 shows the forest plot of differences in effect sizes between trials with and trials without exclusions across the 14 meta-analyses. On average, treatment effects were more beneficial in trials with exclusions than in trials without exclusions (difference in effect sizes −0.13, 95% confidence interval −0.29 to 0.04, P=0.13), but the variability in bias between meta-analyses was considerable (τ^2=0.07, P<0.001). Differences in effect sizes ranged from −0.82 to 0.35. Tests of interaction between exclusions from the analysis and estimates of treatment effects were positive at the conventional level of P=0.05 in five meta-analyses: in four meta-analyses estimated effects were more beneficial in trials with exclusions from the analysis and in one meta-analysis estimated effects were more beneficial in trials without exclusions (fig 2).

Figure 3 presents results of stratified analyses. Differences between trials with and trials without exclusions were evident in meta-analyses with a high degree of between trial heterogeneity, but not in meta-analyses with low between trial heterogeneity (P for interaction <0.001). Similarly, differences were more pronounced in meta-analyses with large estimated treatment benefits in the overall meta-analysis compared with meta-analyses with small estimated benefits (P for interaction <0.001) and in meta-analyses of complementary interventions compared with conventional medicine (P for interaction <0.001). When stratifying for these characteristics, the variability in bias decreased considerably. For example, τ^2 was 0.03 in meta-analyses of complementary medicine and 0.02 in meta-analyses of conventional medicine. When adjusting for concealment of allocation (−0.11, 95% confidence interval

−0.28 to 0.05, P=0.18) or patient blinding (−0.15, −0.30 to 0.00, P=0.047), average differences between trials with and trials without exclusions of patients were robust. In both adjusted analyses the variability in bias between meta-analyses was much the same as in the primary analysis, with variance estimates τ^2 of 0.08 (P<0.001) and 0.06 (P<0.001), respectively.

Impact of restricting meta-analyses to trials without exclusions

Figure 4 presents comparisons of overall meta-analyses including all trials with restricted meta-analyses including trials without exclusions only. After the restriction the number of trials included in a single meta-analysis decreased from a median of 11 to a median of 2 and the number of patients from a median of 1731 to a median of 401. Estimates of treatment benefits decreased in 10 meta-analyses and increased in four (P=0.10). Between trial heterogeneity decreased in 12 meta-analyses and increased in one (P=0.006). For one meta-analysis only one trial had no exclusions from the analysis, and no between trial heterogeneity could be estimated after the restriction.[30] Precisions of pooled effect size estimates decreased in nine meta-analyses and increased in five (P=0.25). P values became larger in 10 meta-analyses and smaller in four (P=0.016). After the restriction to trials without exclusions only, six meta-analyses lost statistical significance at the conventional level of P=0.05.

DISCUSSION

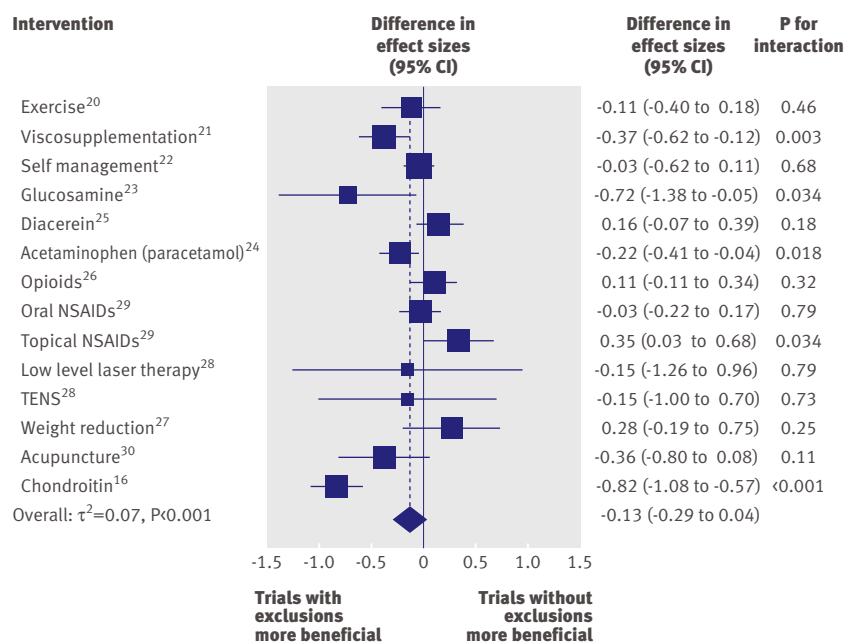In this meta-epidemiological study of 14 meta-analyses and 167 trials we found that excluding randomised patients from the analysis often resulted in biased estimates of treatment effects. The average estimate of bias of a difference in effect size of −0.13 may seem small (fig 2), however it corresponds to one quarter to one half of a typical treatment effect found for interventions in osteoarthritis.[15] The impact of exclusions on estimates of treatment effects seemed most pronounced in meta-analyses with large treatment benefits, meta-analyses on complementary interventions, and meta-analyses with a high degree of heterogeneity between trials, but the extent and direction of bias may be unpredictable in a specific situation. Tests of interaction between exclusions from the analysis and estimates of treatment effects were statistically significant in five meta-analyses; in four of these meta-analyses, estimated treatment effects were less beneficial in trials without exclusions.

When restricting meta-analyses to trials without exclusions, P values increased in most cases and six meta-analyses lost statistical significance at P=0.05 (fig 4). This increase in P values was not only due to a loss of statistical power.[31] As a result of the restriction the between trial heterogeneity τ^2 decreased considerably. Therefore the average loss of statistical precision of random effects meta-analyses was smaller than what could be expected after the exclusion of over half the trials. Only in five meta-analyses was there a relevant loss of precision after the restriction, in six meta-analyses the statistical precision remained much the same, and in three meta-analyses the precision increased.
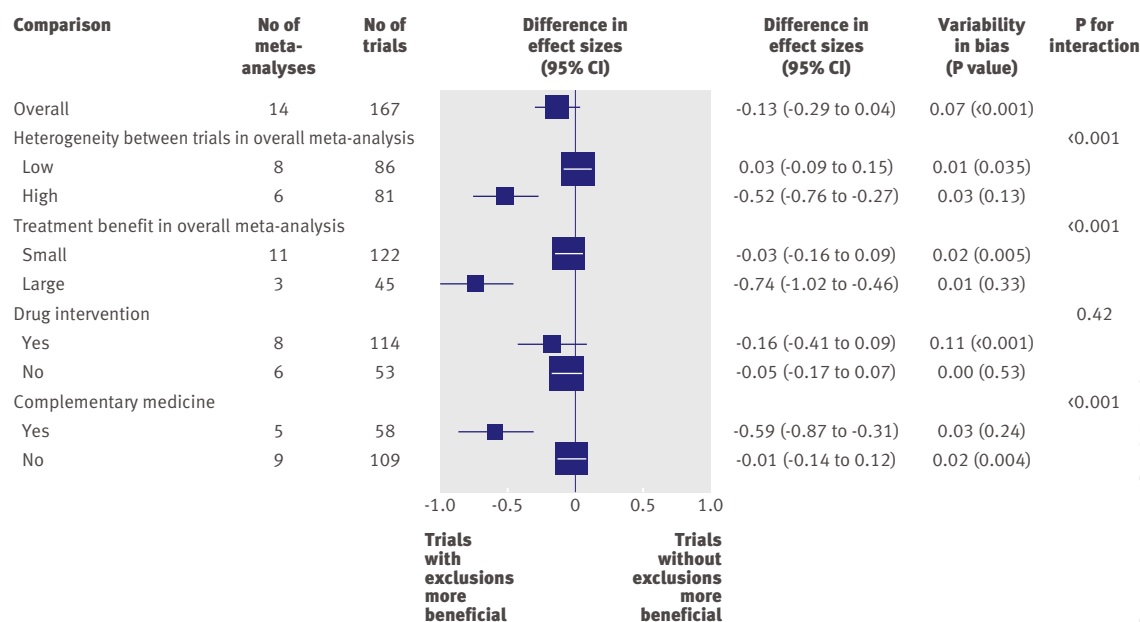
Strengths and limitations of the study

In practice, various definitions of the intention to treat principle are used.[5 32] In our study we did not rely on statements in the trial reports on whether an intention to treat analysis was done or not. Rather we required explicit information about the flow of patients through the study[33 34] or clear statements that all randomised patients were included in the analysis. Some might argue that our distinction between trials with and trials without exclusions from the analysis was overly stringent. The exclusion of only a small proportion of patients from the analysis, for example, may be considered unlikely to have any impact on estimated treatment benefits. We would expect that any bias associated with exclusions from the analysis will increase with the number of exclusions. Therefore the overall estimate of bias might increase with the selection of a less rigorous cut off. Others may argue that our classification was not stringent enough and that we should have required an affirmative statement that no crossovers had occurred and that all randomised patients were included in the analysis in the group to which they were originally allocated. Only seven of the 167 included trials (4%) explicitly provided this information, so we were unable to examine this issue.

As with other meta-epidemiological studies,[6 9] our study is based on published information and depends on the quality of the trial reports. Even though the quality of reporting is generally low,[35 36] we were able to

Intervention	Difference in effect sizes (95% CI)	Difference in effect sizes (95% CI)	P for interaction
Exercise[20]		−0.11 (−0.40 to 0.18)	0.46
Viscosupplementation[21]		−0.37 (−0.62 to −0.12)	0.003
Self management[22]		−0.03 (−0.62 to 0.11)	0.68
Glucosamine[23]		−0.72 (−1.38 to −0.05)	0.034
Diacerein[25]		0.16 (−0.07 to 0.39)	0.18
Acetaminophen (paracetamol)[24]		−0.22 (−0.41 to −0.04)	0.018
Opioids[26]		0.11 (−0.11 to 0.34)	0.32
Oral NSAIDs[29]		−0.03 (−0.22 to 0.17)	0.79
Topical NSAIDs[29]		0.35 (0.03 to 0.68)	0.034
Low level laser therapy[28]		−0.15 (−1.26 to 0.96)	0.79
TENS[28]		−0.15 (−1.00 to 0.70)	0.73
Weight reduction[27]		0.28 (−0.19 to 0.75)	0.25
Acupuncture[30]		−0.36 (−0.80 to 0.08)	0.11
Chondroitin[16]		−0.82 (−1.08 to −0.57)	<0.001
Overall: τ^2=0.07, P<0.001		−0.13 (−0.29 to 0.04)	

-1.5 -1.0 -0.5 0 0.5 1.0 1.5

Trials with exclusions more beneficial | Trials without exclusions more beneficial

Fig 2 | Difference in effect sizes between 128 trials with and 39 trials without exclusions of patients from analysis. A negative difference in effect sizes indicates that trials with exclusions of patients from analysis show more beneficial treatment effects. P values are for interaction between exclusions from analysis and effect sizes. NSAIDs=non-steroidal anti-inflammatory drugs; TENS=transcutaneous electrical nerve stimulation

Comparison	No of meta-analyses	No of trials	Difference in effect sizes (95% CI)	Difference in effect sizes (95% CI)	Variability in bias (P value)	P for interaction
Overall	14	167		-0.13 (-0.29 to 0.04)	0.07 (<0.001)	
Heterogeneity between trials in overall meta-analysis						<0.001
Low	8	86		0.03 (-0.09 to 0.15)	0.01 (0.035)	
High	6	81		-0.52 (-0.76 to -0.27)	0.03 (0.13)	
Treatment benefit in overall meta-analysis						<0.001
Small	11	122		-0.03 (-0.16 to 0.09)	0.02 (0.005)	
Large	3	45		-0.74 (-1.02 to -0.46)	0.01 (0.33)	
Drug intervention						0.42
Yes	8	114		-0.16 (-0.41 to 0.09)	0.11 (<0.001)	
No	6	53		-0.05 (-0.17 to 0.07)	0.00 (0.53)	
Complementary medicine						<0.001
Yes	5	58		-0.59 (-0.87 to -0.31)	0.03 (0.24)	
No	9	109		-0.01 (-0.14 to 0.12)	0.02 (0.004)	

-1.0 -0.5 0 0.5 1.0

Trials with exclusions more beneficial — Trials without exclusions more beneficial

Fig 3 | Differences in effect sizes between 128 trials with and 39 trials without exclusions of patients from analysis stratified according to four characteristics of meta-analyses. See table 1 for a description of meta-analyses according to these characteristics. A $\tau^2 < 0.06$ indicates low between trial heterogeneity and a $\tau^2 \geq 0.06$ high between trial heterogeneity. An effect size > -0.5 indicates a small benefit of the experimental intervention and an effect size ≤ -0.5 a large benefit. Meta-analyses are ordered according to year of publication. A negative difference in effect sizes indicates that trials with exclusions of patients from analysis show a more beneficial treatment effect. Variability in bias between-meta-analyses is expressed as heterogeneity variance τ^2

determine in all but 14 trials whether exclusions from the analysis had occurred. Compared with previous meta-epidemiological studies,[1 6 8-10] misclassification of trials due to inadequate reporting[11] is therefore less likely to have introduced bias in our study. At least two thirds of the trials included in our study had incomplete outcome data. Two approaches towards imputation of missing data are generally used to replace missing data and allow an intention to treat analysis: the last observation carried forward method and multiple imputation. We were unable to examine whether the approach used for data imputation influences estimates of treatment effects because of the strikingly low quality of reporting.[32 34 35] Other types of bias that may affect the results of randomised trials include selection bias due to inadequate concealment of allocation, and performance and assessor bias due to a lack of blinding of patients and therapists.[2 12] In our study, the observed association between exclusions of patients from the analysis and estimates of treatment effects could be confounded by concealment of allocation: trials with exclusions tended to report adequate concealment of allocation less often than trials without exclusions. This correlation may have resulted in spurious associations. When accounting for concealment of allocation in a sensitivity analysis, however, we found our results to be robust. Finally, characteristics of meta-analyses were also correlated. For example, meta-analyses in complementary medicine were likely to show large treatment benefits and a high degree of heterogeneity between trials. Our understanding of the interplay of these characteristics is incomplete. Therefore the

results of our stratified analyses (fig 3) should to be interpreted with caution. A detailed examination of that problem would require a much larger set of meta-analyses.

Context

As is the case for other types of bias,[12] the extent of attrition bias might depend on the type of outcome. Ours is the first meta-epidemiological study to investigate pain as a patient reported outcome, a measure extensively used in research on osteoarthritis.[15] This outcome is more prone to bias than objective binary outcomes such as mortality.[12] Four studies have examined the impact of attrition bias on estimates of treatment benefits in randomised controlled trials and meta-analyses on an odds ratio scale.[1 6 8-10] The direction and magnitude of attrition bias varied between different studies according to different methods and definitions used, different clinical areas addressed, and the potential for misclassification of trials because of inadequate reporting.[6 8-10] A recent study used individual patient data and found that analyses with patients excluded showed more beneficial effects of the experimental treatment than analyses according to the intention to treat principle.[1] Another study of placebo controlled trials of serotonin reuptake inhibitors sponsored by the pharmaceutical industry found that the experimental intervention was favoured less in intention to treat analyses than in per protocol analyses. Many published reports of these trials ignored the results of intention to treat analyses and reported only the more favourable per protocol analyses.[37]
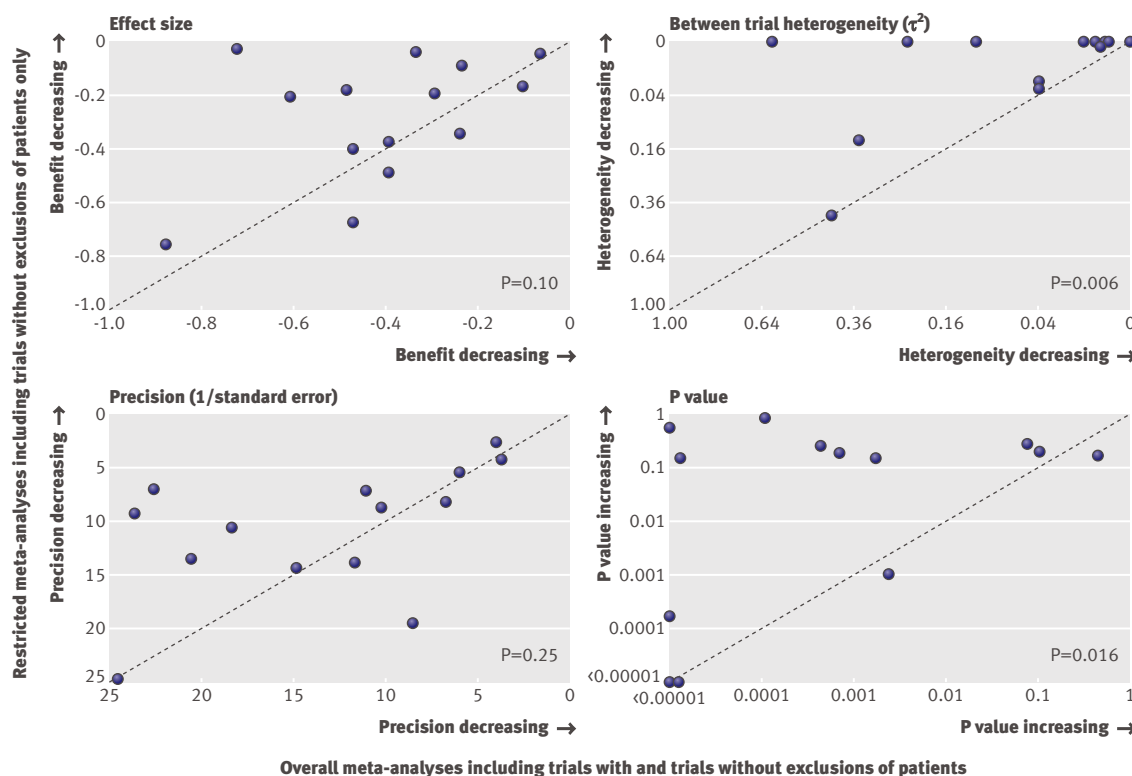
Fig 4 | Effect sizes, between trial heterogeneity τ^2, precision, and P values of overall treatment benefits compared between overall meta-analyses including trials with and without exclusions of patients (x axis) and restricted meta-analyses including trials without exclusions of patients only (y axis). Dashed line indicates that estimates are identical. P values are derived from Wilcoxon rank tests for paired observations

Several authors pointed out that attrition bias can go in either direction and is difficult to predict for a specific situation,[1 2] which is in accordance with our findings of highly variable effects between meta-analyses. Previous meta-epidemiological studies, which examined the effect of exclusions from the analysis,[1 6 8-10] might be reanalysed in the light of our results to examine the variability in bias associated with exclusions.

Implications

The intention to treat principle aims to compare patients in the groups to which they were originally allocated. The most stringent interpretation of intention to treat includes the analysis of all patients, regardless of whether they were eligible, received treatment, and adhered to the protocol.[5] In practice, various interpretations are used, some of which allow for exclusions after randomisation. Many trialists exclude randomised patients who did not receive at least one dose of the allocated intervention, whereas others exclude patients found retrospectively to be ineligible.[5 38] Both approaches to excluding patients from the analysis may produce unbiased estimates if patients and treating doctors are unaware of the allocated intervention and if the decision to exclude patients is based solely on information collected before randomisation and unrelated to group assignment and clinical outcome.[38] In addition, exclusions from the analysis owing to randomly missing outcome data may be less problematic

than the selective exclusion of patients owing to protocol violations. These assumptions are hardly ever verifiable: details on the flow of participants through the various stages of a trial and descriptions of procedures used to determine whether patients should be excluded from the analysis are often omitted from published reports of randomised trials.[5 34] Therefore it is difficult to determine from published information whether reported exclusions from the analyses resulted in bias,[2] and strict adherence to the intention to treat principle should be advocated.[33 39]

The purpose of an intention to treat analysis is to preserve an unbiased treatment allocation and the prognostic balance between treatment groups. In contrast, per protocol analyses include only those patients who received treatment as defined in the study protocol and provided outcome data. Patients excluded from per protocol analyses are likely to be different from those analysed: they may have had an acute exacerbation of the studied condition or experienced side effects of the evaluated intervention.[3] Trials without exclusions more often reported imputations of missing data than those with exclusions. The last observation carried forward approach was used most often: missing values were replaced by the last value observed. This method is popular for imputation of missing data in musculoskeletal research[40 41] but leads to overly precise estimates and potential bias.[42 43] Multiple imputation is more difficult to carry out but avoids

---

Recommendations for practice

- Since excluding patients from the analysis often results in biased estimates of treatment effects, trialists should ensure low dropout rates and high compliance rates and minimise missing outcome data

- Trialists should always report results of intention to treat analyses, including all randomised patients in the analysis in the group to which they were originally allocated. If data imputations are necessary to carry out an intention to treat analysis, multiple imputation should be used to replace missing data

- Those critically appraising trials should generally rely on results from intention to treat analyses

- Authors of systematic reviews should routinely examine the influence of exclusions from the analysis on estimated treatment effects. In case of discrepancies between trials with and trials without exclusions, trials without exclusions should be given precedence

those problems[44]: each missing value is replaced by multiple simulated values, and the analysis of the resulting multiple versions of the complete dataset can account for the uncertainty about missing data. The CONSORT statement urges transparent reporting of the flow of participants through the various stages of a trial, including a description of withdrawals and losses to follow-up and the reasons for exclusions from the analysis.[33 39] In our view a detailed description of strategies used to handle missing outcome data is also essential.

Conclusions

The box summarises our recommendations for practice. Excluding patients from the analysis often results in biased estimates of treatment effects in randomised trials. To avoid potential attrition bias, trialists should ensure low dropout rates and high compliance rates and minimise missing outcome data. Results of intention to treat analyses, which are based on the inclusion of all patients in the analysis in the group to which they were originally allocated, should always be reported. Sensitivity analyses, which are restricted to patients adhering to the protocol, may be described in addition. In systematic reviews and meta-analyses, data

extraction should be based on results from analyses of all randomised patients, whenever possible. The influence of exclusions from analysis on estimated treatment benefits should be routinely assessed in stratified analyses. This may be particularly important in complementary medicine, in the presence of high heterogeneity between trials, and when pooled effect sizes indicate a large benefit of evaluated interventions.

1 Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005;34:79-87.
2 Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
3 Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;301:1410-2.
4 Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *N Engl J Med* 1980;303:1038-41.
5 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670-4.
6 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.
7 Sterne JA, Juni 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.
8 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.
9 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.
10 Gluud LL, Thorlund K, Gluud C, Woods L, Harris R, Sterne JA. Correction: reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2008;149:219.
11 Juni P, Egger M. Commentary: empirical evidence of attrition bias in clinical trials. *Int J Epidemiol* 2005;34:87-8.
12 Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Excluding randomised patients from the analysis in randomised controlled trials may result in attrition bias

The direction and magnitude of attrition bias varies between different studies according to different methods and definitions used and different clinical areas addressed

WHAT THIS STUDY ADDS

Excluding patients from the analysis of randomised trials often resulted in biased estimates of treatment effects, but the extent and direction of bias remained unpredictable in a specific situation

Overestimation of treatment benefits seemed particularly pronounced in meta-analyses with high between trial heterogeneity, in meta-analyses with large estimated treatment benefits, and in meta-analyses of complementary medicine

Restriction of meta-analyses to trials without exclusions of patients resulted in smaller estimated treatment benefits, larger P values, and considerable decreases in between trial heterogeneity

13 Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005;330:68.

14 Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002;31:140-9.

15 Juni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol* 2006;20:721-40.

16 Reichenbach S, Sterchi R, Scherer M, Trelle S, Burgi E, Burgi U, et al. Meta-analysis: chondroitin for osteoarthritis of the knee or hip. *Ann Intern Med* 2007;146:580-90.

17 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.

18 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Earlbaum, 1988.

19 Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian approaches to clinical trials and health-care evaluation*. Chichester: Wiley, 2004.

20 Fransen M, McConnell S, Bell M. Exercise for osteoarthritis of the hip or knee. *Cochrane Database Syst Rev* 2003(3):CD004286.

21 Lo GH, LaValley M, McAlindon T, Felson DT. Intra-articular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis. *JAMA* 2003;290:3115-21.

22 Chodosh J, Morton SC, Mojica W, Maglione M, Suttorp MJ, Hilton L, et al. Meta-analysis: chronic disease self-management programs for older adults. *Ann Intern Med* 2005;143:427-38.

23 Towheed TE, Maxwell L, Anastassiades TP, Shea B, Houpt J, Robinson V, et al. Glucosamine therapy for treating osteoarthritis. *Cochrane Database Syst Rev* 2005(2):CD002946.

24 Towheed TE, Maxwell L, Judd MG, Catton M, Hochberg MC, Wells G. Acetaminophen for osteoarthritis. *Cochrane Database Syst Rev* 2006(1):CD004257.

25 Rintelen B, Neumann K, Leeb BF. A meta-analysis of controlled clinical studies with diacerein in the treatment of osteoarthritis. *Arch Intern Med* 2006;166:1899-906.

26 Avouac J, Gossec L, Dougados M. Efficacy and safety of opioids for osteoarthritis: a meta-analysis of randomized controlled trials. *Osteoarthritis Cartilage* 2007;15:957-65.

27 Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. *Ann Rheum Dis* 2007;66:433-9.

28 Bjordal JM, Johnson MI, Lopes-Martins RA, Bogen B, Chow R, Ljunggren AE. Short-term efficacy of physical interventions in osteoarthritic knee pain. A systematic review and meta-analysis of randomised placebo-controlled trials. *BMC Musculoskelet Disord* 2007;8:51.

29 Bjordal JM, Klovning A, Ljunggren AE, Slordal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: a meta-analysis of randomised placebo-controlled trials. *Eur J Pain* 2007;11:125-38.

30 Manheimer E, Linde K, Lao L, Bouter LM, Berman BM. Meta-analysis: acupuncture for osteoarthritis of the knee. *Ann Intern Med* 2007;146:868-77.

31 Pildal J, Hrobjartsson A, Jorgensen K, Hilden J, Altman D, Gotzsche P. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007;36:847-57.

32 Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? *Clin Trials* 2007;4:350-6.

33 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.

34 Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996-9.

35 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.

36 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.

37 Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine—selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;326:1171-3.

38 Fergusson D, Aaron SD, Guyatt G, Hébert P. Postrandomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002;325:652-4.

39 Moher D, Schulz KF, Altman DG. THE CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001;134:657-62.

40 Baron G, Boutron I, Giraudeau B, Ravaud P. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. *Arthritis Rheum* 2005;52:1858-65.

41 Kim M. Statistical methods in Arthritis & Rheumatism: current trends. *Arthritis Rheum* 2006;54:3741-9.

42 Julious SA, Mullee MA. Issues with using baseline in last observation carried forward analysis. *Pharm Stat* 2008;7:142-6.

43 Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. *Can J Psychiatry* 2002;47:68-75.

44 Baron G, Ravaud P, Samson A, Giraudeau B. Missing data in randomized controlled trials of rheumatoid arthritis with radiographic outcomes: a simulation study. *Arthritis Rheum* 2008;59:25-31.