

REVIEW ARTICLE

No differential attrition was found in randomized controlled trials published in general medical journals: a meta-analysis

Rik Crutzen^{a,*}, Wolfgang Viechtbauer^b, Daniel Kotz^{a,c}, Mark Spigt^{a,d}

^aCAPHRI School for Public Health and Primary Care, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

^bMHeNS School for Mental Health and Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

^cDepartment of Epidemiology and Public Health, Cancer Research UK Health Behaviour Research Centre, University College London, WC1E 6BT, UK

^dDepartment of General Practice, Tromsø University, Norway

Accepted 12 March 2013; Published online 21 June 2013

Abstract

Objective: Differential attrition is regarded as a major threat to the internal validity of a randomized controlled trial (RCT). This study identifies the degree of differential attrition in RCTs covering a broad spectrum of clinical areas and factors that are related to this.

Study Design and Setting: A PubMed search was conducted to obtain a random sample of 100 RCTs published between 2008 and 2010 in journals from the ISI Web of KnowledgeSM category of medicine, general and internal. Eligibility criteria for selecting studies were primary publications of two-arm parallel randomized clinical trials, containing human participants and one or multiple follow-up measurements whose availability depended on the patients' willingness to participate.

Results: A significant amount of differential attrition was observed in 8% of the trials. The average differential attrition rate was 0.99 (95% confidence interval: 0.97–1.01), indicating no general difference in attrition rates between intervention and control groups. Moreover, no indication of heterogeneity was found, suggesting that the occurrence of differential attrition in the published literature is mostly a chance finding, unrelated to any particular design factors.

Conclusion: Differential attrition did not generally occur in RCTs covering a broad spectrum of clinical areas within general and internal medicine. © 2013 Elsevier Inc. All rights reserved.

Keywords: Differential attrition; RCT; Internal validity; Meta-analysis; Bias; Loss to follow-up

1. Introduction

Attrition or loss to follow-up after randomization is a common problem in randomized controlled trials (RCTs) [1,2], which complicates the statistical analyses and can lead to bias in the findings [3]. When the degree of attrition differs between the various treatment groups that are being compared in an RCT, then this is typically called differential attrition. Because these groups are, if random allocation is undertaken properly, comparable at baseline (e.g., in terms of the study's primary outcome), differential attrition can be assumed to be a consequence of differences between the groups that arose at some point after randomization (e.g., because of perceived treatment efficacy, safety, or tolerability) [4]. For example, patients in the control group who improve in terms of the study's primary outcome

might be more prone to complete follow-up measures than those who do not improve, or treated patients (i.e., those in the intervention group) may feel a general obligation to complete follow-up measures [5]. Therefore, differential attrition is usually regarded as a major threat to the internal validity of a study (i.e., whether the intervention really did cause a change in the outcome) [6], but insight into the degree of differential attrition occurring in RCTs is limited.

Previous meta-analyses focusing on a single clinical area found no differential attrition in trials regarding interventions aimed at self-monitoring of blood glucose in type 2 diabetes [7] and the use of serotonin-specific reuptake inhibitors in treatment of posttraumatic stress disorder [8] but did find differences when comparing atypical and typical antipsychotic medications [9]. In a convenience sample of 10 trials evaluating interventions for the treatment of musculoskeletal disorders, all trials showed some level of differential attrition between the treatment arms, ranging from 1% to 14% [10]. A systematic review of comprehensive cohorts and two-stage trials that measured or recorded patient or physician

Conflict of interest/Financial disclosure: Neither conflict of interest nor financial support regarding the present study was reported.

* Corresponding author. Tel.: +31-43-3882828; fax: +31-43-3671032.
E-mail address: Rik.Crutzen@maastrichtuniversity.nl (R. Crutzen).

What is new?

- This study did not find evidence that differential attrition generally occurs in RCTs covering a broad spectrum of clinical areas.
- Potentially anticipatable sources of differential attrition (e.g., differences in the number of contact moments, additional advantages because of taking part in the intervention) should be taken into account when designing an RCT.
- The average overall attrition is 13%, which can be used as a rule of thumb in sample size calculations of future RCTs within general and internal medicine.

preference found little evidence that preference affects validity [11]. Furthermore, Puffer et al. [12] found evidence of differential attrition in 4 of 36 cluster randomized trials published in three general medicine journals. In only one of these trials did the authors comment on this in the discussion section.

To our knowledge, no study exists that has examined differential attrition in general (i.e., not focused on a single clinical area). This makes it impossible to draw any general conclusions about the degree of differential attrition in RCTs. We therefore conducted the present study to address the following key question: How often and to what degree does differential attrition occur in RCTs covering a broad spectrum of clinical areas within general and internal medicine, and which factors may be related to the degree to which differential attrition occurs in such studies?

2. Methods

A systematic review and meta-analysis of a random selection of RCTs published between 2008 and 2010 in journals from the ISI Web of KnowledgeSM category of medicine, general and internal was conducted.

2.1. Search strategy

To identify the trials, we obtained the impact factors (IFs) from the Journal Citation Reports Science Edition 2009 of ISI Web of KnowledgeSM. To obtain a representative sample of articles from both higher and lower impact journals, we selected the top 40 journals (IF: mean, 6.8; standard deviation [SD], 9.3; median, 2.9; range, 1.8–47.1), ranked these journals based on their IF, and created 4 groups of 10 journals, group 1 being the group with the highest IF and group 4 with the lowest. Subsequently, we searched PubMed for articles in each group with the

following restrictions: (“Randomized Controlled Trial” [PT] OR “Controlled Clinical Trial” [PT] OR “Clinical Trial” [PT]) AND humans [MeSH Terms] AND English [LA] AND 2008:2010 [DP] AND random* [TIAB]). This resulted in 1,150 hits for group 1; 221 hits for group 2; 504 hits for group 3; and 99 hits for group 4; a total of 1,974.

2.2. Selection of trials

Each study identified in this manner then received a trial identification number and was randomly assigned to one of the investigators, who examined whether the article fitted the inclusion criteria. In particular, the article had to be the primary publication of a randomized clinical trial with an identifiable intervention and control/comparison group. Trials with more than two treatment groups or multifactor studies were included if groups could be collapsed in a logical manner (e.g., groups receiving varying dosages of a medication could be collapsed into a single intervention group). Furthermore, the trial should contain human participants and one or multiple follow-up measurements whose availability depended on the patients’ willingness to participate (e.g., a trial that measured outcomes while patients stayed in the clinic was excluded because of the limited chance of loss to follow-up). If in doubt, inclusion of the article was discussed with the rest of the research team until a unanimous decision was reached. If an article did not fit the inclusion criteria, a new article was randomly selected based on the screening list.

This process was continued until the desired number of studies was selected from each group. Proportional stratified sampling (i.e., in proportion to the number of hits per group) was used to select articles, with the goal of obtaining approximately 5% of the trials within each group. Accordingly, we included articles until the following number of articles from each group was selected: 58 articles from group 1; 12 articles from group 2; 24 articles from group 3; and 6 articles from group 4; resulting in a total of 100 articles (Fig. 1).

2.3. Analysis of RCTs

The final set of 100 articles included a wide variety of trials, covering the following general categories: various types of medication for a variety of different medical conditions (e.g., uncomplicated acute cystitis, superficial vein thrombosis), the benefits of vitamin supplementation (e.g., vitamin K in postmenopausal women with osteopenia), surgical procedures (e.g., male circumcision, stent placement), vaccination (e.g., hepatitis B), behavioral/counseling interventions (e.g., computerized tailored physical activity reports), and policy evaluations (e.g., removing direct payment for health care). Each included RCT was independently scored on the items shown in Table 1 by two investigators (the one who initially examined the article for

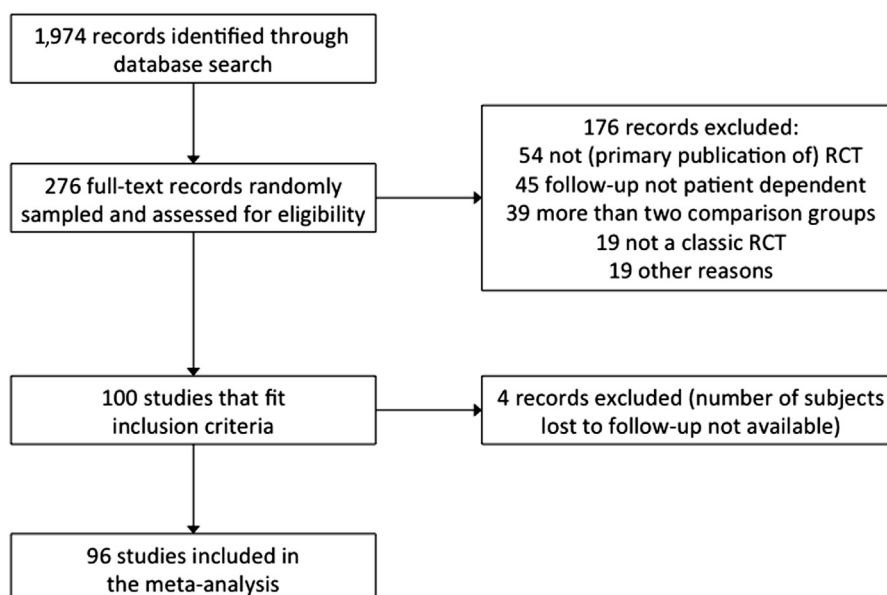


Fig. 1. Flowchart of study selection. RCT, randomized controlled trial.

suitability and one additional randomly chosen investigator). Besides descriptive characteristics, these items concerned factors potentially related to differential attrition (e.g., intensity of treatment per group, type of control; Table 1). These items were selected based on their face validity in line with the study aim, determined on by all investigators before the start of collecting data. Disagreements between the two scorers were resolved by discussion. The intensity of the intervention, control condition, and follow-up procedure were scored on a scale from 1 (lowest intensity; e.g., only an information leaflet or single questionnaire) to 10 (highest intensity; e.g., severe surgical intervention or long-term diary registration). The number of participants who remained in the study in the intervention and control groups was determined based on the primary outcome indicated by the authors. Moreover, when multiple follow-up measurements were taken, we recorded the number that remained at the last time point that contributed to the primary analysis (e.g., at 12 months, for a study with measurements at baseline and at 3, 6, 9, 12, and 24 months after treatment, where the primary analysis is a repeated-measures analysis for the data up to month 12). For each study, we therefore obtained a 2×2 table with the number of participants who were lost to follow-up and the number that remained in the study in the intervention and control groups.

2.4. Statistical analysis

For each study, we determined the proportion of participants who were lost to follow-up after randomization in the intervention and control groups (p_T and p_C , respectively) based on the 2×2 table. We then calculated the relative attrition rate (i.e., p_T/p_C), so that a relative attrition

rate greater than one indicates a higher attrition rate in the intervention group, whereas a rate less than one indicates a higher attrition rate in the control group. Studies with no dropout in either the intervention or the control group were handled by adding half to each cell in the 2×2 table, which not only makes the computation of the log relative attrition rate possible but also more generally acts as a bias correction [13]. Studies with no dropout in both groups were excluded from the analysis (a sensitivity analysis including such studies leads to unchanged conclusions) [13].

The log-transformed relative attrition rates were then meta-analyzed based on a random-effects model, using restricted maximum likelihood (REML) estimation to estimate the amount of heterogeneity. We reported the (back-transformed) estimated average relative attrition rate and corresponding 95% confidence interval (CI), in addition to the results from the Q test for heterogeneity and the I^2 statistic. The rank correlation and regression test were used to examine whether asymmetry was present in the funnel plot [14].

Meta-regression analyses using mixed-effects models were then conducted to examine whether the degree of differential attrition was related to the IF group, difference between the treatment intensity in the intervention and control groups, type of control group, whether participants were blinded or not, follow-up duration, and follow-up intensity. We examined these potential moderators univariately and also jointly in a single meta-regression model. Again, REML estimation was used to fit these models.

For all models, we checked the data for outliers and influential cases using residuals and various case and deletion

Table 1. Characteristics of included studies, stratified by group of journals

Characteristics	Group				Total (N = 100)
	1 (N = 58)	2 (N = 12)	3 (N = 24)	4 (N = 6)	
Intensity of the intervention ^a	4.5 (1.8)	3.8 (2.1)	3.9 (1.5)	4.8 (2.1)	4.3 (1.8)
Intensity of the control ^a	3.5 (1.8)	2.4 (2.0)	3.7 (1.8)	4.8 (2.1)	3.5 (1.9)
Type of control, % (N)					
Wait list/nothing	5.2 (3)	25.0 (3)	4.2 (1)	0	7.0 (7)
Care as usual	20.7 (12)	33.3 (4)	12.5 (3)	0	19.0 (19)
Placebo/sham intervention	46.6 (27)	8.3 (1)	37.5 (9)	33.3 (2)	39.0 (39)
Alternative intervention	27.6 (16)	33.3 (4)	45.8 (11)	66.7 (4)	35.0 (35)
Blinding of participants (yes), % (N)	18.8 (3/16)	0	54.5 (6/11)	25.0 (1/4)	28.6 (10/35)
Follow-up duration (mo), median (IQR)	10.0 (15.0)	6.5 (9.0)	2.0 (5.2)	1.4 (2.4)	6.0 (10.0)
Follow-up intensity ^a	3.4 (1.9)	2.7 (1.0)	3.6 (1.6)	4.5 (2.4)	3.4 (1.8)
Intention-to-treat analysis mentioned (yes), % (N)	86.2 (50)	50.0 (6)	58.3 (14)	33.3 (2)	72.0 (72)
Participants randomized into intervention, median (IQR)	167.0 (404)	182.5 (322)	62.5 (145)	44.0 (41)	120.5 (235)
Participants randomized into control, median (IQR)	158.0 (402)	148.0 (240)	73.5 (129)	44.5 (40)	119.0 (217)
Participants retained in intervention, median (IQR)	137.0 (316)	159.5 (247)	42.0 (140)	38.0 (43)	107.5 (194)
Participants retained in control, median (IQR)	117.0 (317)	108.5 (194)	55.0 (121)	35.0 (40)	103.5 (186)
Handling of attrition, % (N)					
Nothing (complete case analysis)	22.4 (13)	41.7 (5)	41.7 (10)	33.3 (2)	30.0 (30)
Imputation of missing values	37.9 (22)	41.7 (5)	25.0 (6)	33.3 (2)	35.0 (35)
All available data analyzed	27.6 (16)	8.3 (1)	12.5 (3)	0	20.0 (20)
Other method	0	0	4.2 (1)	0	1.0 (1)
Unknown	8.6 (5)	8.3 (1)	12.5 (3)	33.3 (2)	11.0 (11)
Attrition discussed in article (yes), % (N)	46.6 (27)	66.7 (8)	20.8 (5)	0	40.0 (40)
If attrition discussed, differential attrition discussed in article (yes), % ^b (N)	55.6 (15/27)	37.5 (3/8)	80.0 (4/5)	NA	55.0 (22/40)

Abbreviations: IQR, interquartile range; NA, not available.

Data presented as mean (standard deviation), unless stated otherwise. Groups of journals defined on the basis of 2009 impact factors.

^a Intensity measured on a scale from 1 (lowest) to 10 (highest intensity).

^b Percentage of articles in which attrition was discussed (N = 40).

diagnostics [15]. All analyses were conducted in R [16] using the *metafor* package (written by W.V.) [17].

3. Results

Of the 1,974 articles that were obtained from the PubMed search, a total of 276 articles had to be screened to identify 100 eligible articles (36%). The numbers of articles per group (and total number of articles screened for eligibility in that group) were 58 (131) in group 1; 12 (32) in group 2; 24 (93) in group 3; and 6 (20) in group 4. Therefore, a total of 176 articles were not eligible and excluded from the analyses: 54 (31%) because the article did not present data from an RCT or because the article was not the primary publication of the trial; 45 (26%) because the intervention or the follow-up measurement was not patient dependent; 39 (22%) because the trial had more than two comparison groups (and collapsing of groups was not possible); 19 (11%) because the study was not a classic RCT (e.g., Zelen design or equivalence trial), and 19 (11%) for other reasons (e.g., “zero” follow-up time, or in other words, the outcome of a study was assessed during the same session as the treatment).

For 4 of the 100 studies (three in IF group 1 and one in IF group 4), we were unable to extract or determine the

number of subjects that were lost to follow-up in the intervention and control groups. In the remaining 96 studies, the attrition rates ranged from 0 to 0.54 (0–54%) in the intervention and from 0 to 0.62 (0–62%) in the control groups. The average attrition rate was 0.13 (SD, 0.11; median, 0.11) in both groups. As shown in Fig. 2, the distribution of the attrition rates was heavily right skewed in both groups.

In 6 of the 96 studies, there was no dropout in both the intervention and control groups. These studies were therefore removed from the analysis. In the remaining 90 studies, the relative attrition rates ranged from 0.19 to 5.00, with the distribution centered visibly at a relative rate of 1 (Fig. 3). The distribution was not entirely balanced: in 53 of 89 cases (60%), the attrition rate was larger in the control group (leaving out one study with a relative attrition rate exactly equal to 1).

In 7 of the 90 studies (8%), the relative attrition rate differed significantly from 1 at $\alpha = 0.05$ (two sided), which is quite close to the number of significant findings one would expect to obtain if the true relative attrition rate was equal to 1 in all the studies. In fact, the average relative attrition rate as estimated by the random-effects model was essentially indistinguishable from 1 (i.e., 0.99 with 95% CI: 0.97–1.01). Moreover, the relative attrition rates were not heterogeneous ($Q[\text{df} = 89] = 87.86$, $P = 0.51$; $I^2 = 0\%$), indicating that actual differential attrition observed in any

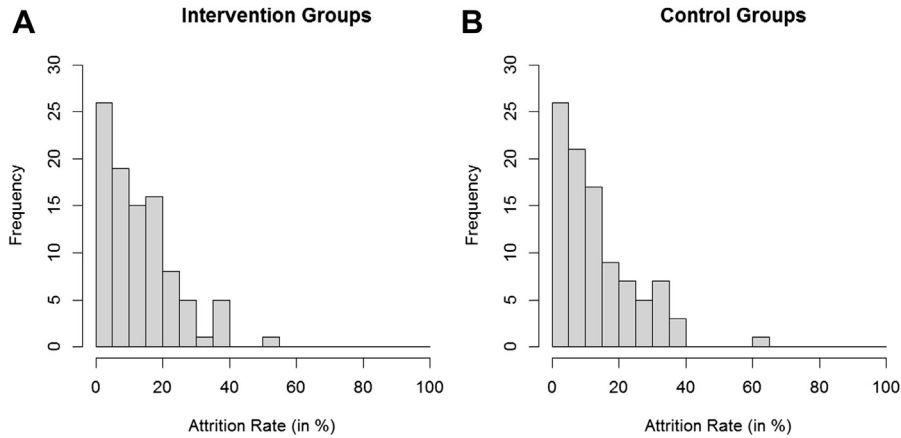


Fig. 2. Distribution of the attrition rates in the (A) intervention and (B) control groups.

particular trial can be attributed to pure random sampling fluctuations. Three studies had disproportionately large weights (one of which was based on more than 100,000 subjects) and therefore were quite influential, but their removal did not change any of the conclusions. The funnel plot for all 90 studies is shown in Fig. 4. Neither the regression test ($P = 0.34$) nor the rank correlation test ($P = 0.55$) indicated asymmetry in the plot.

All the studies were unblinded where the control group consisted of a wait list or received no treatment at all. Moreover, all studies (except one) were unblinded where the control group received care as usual. On the other hand, all the studies using placebos or sham interventions were, by definition, blinded. Only when the control group received an alternative intervention was there any differentiation between blinded and unblinded studies. Therefore, to reduce the number of moderators tested, we collapsed these two moderators into a single moderator with five levels

(Table 2). A mixed-effects model indicated that the average attrition rates differed slightly across these five levels, with studies using alternative intervention control groups with blinding being significantly different from the other types (all pairwise $P \leq 0.023$). In those studies, the attrition rate was estimated to be approximately 35% higher in the control groups. A close look at these studies revealed that this mostly concerned trials comparing two active interventions, in particular, two types of drugs.

There were also some slight differences in the estimated attrition rates across the four IF groups, but none of the pairwise differences was significant (all $P \geq 0.45$). Additional meta-regression models indicated that the relative attrition rate was not influenced by differences in the treatment intensity between the intervention and control groups ($P = 0.51$), follow-up duration ($P = 0.33$), or follow-up intensity ($P = 0.43$). Finally, when using backward elimination with a meta-regression model where all moderators

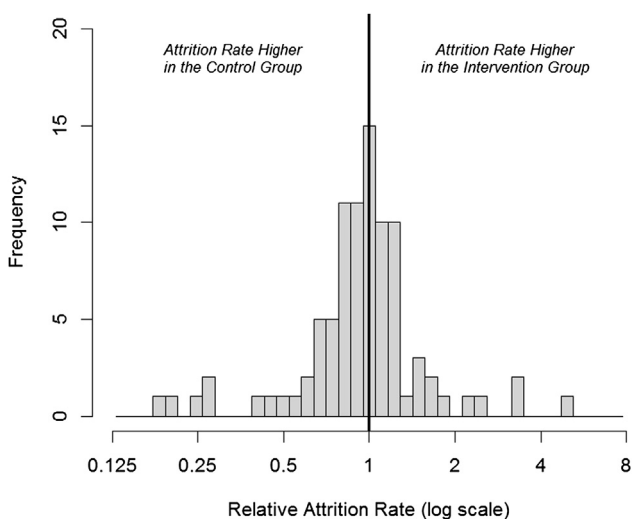


Fig. 3. Distribution of the relative attrition rates.

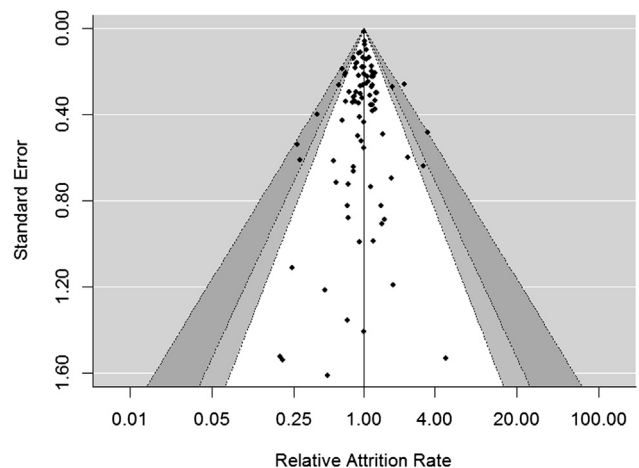


Fig. 4. Contour-enhanced funnel plot of the relative attrition rates (> 1 indicates higher attrition rate in intervention group). White area: $P \geq 0.1$; light gray area: $0.05 \leq P < 0.1$; dark gray area: $0.01 \leq P < 0.05$; area outside funnel: $P < 0.01$.

Table 2. Estimated average relative attrition rate (and 95% CI) for the moderators IF group, type of control, and blinding

Moderator	Levels	Number of studies ^a	Estimate	95% CI	
Type of control	Waiting list/nothing	6	1.05	0.93–1.19	
	Care as usual	19	0.97	0.90–1.05	
	Placebo/sham intervention	32	1.00	0.97–1.02	
	Alternative intervention	Without blinding	24	0.97	0.88–1.07
		With blinding	9	0.74	0.60–0.92
IF group	1	53	0.99	0.96–1.02	
	2	12	0.97	0.89–1.06	
	3	21	1.02	0.91–1.15	
	4	4	0.88	0.55–1.40	

Abbreviations: 95% CI, 95% confidence interval; IF, impact factor.

^a Number of studies included in the meta-regression analyses.

were included initially, we found again that the type of control group was the only significant factor.

As a post hoc analysis, we closely inspected the top 10% and bottom 10% of studies (denoted as “extreme studies”) in terms of the relative attrition rate to identify possible factors related to differential attrition. These extreme studies mostly concerned trials with relatively small sample sizes (n per group < 50), which increase the likelihood of a study being labeled as extreme. In the trials with relatively larger sample sizes, there was more contact with patients in one of the groups in three trials (offering follow-up services by a team [18], differences in time being discharged from the hospital [19], and a self-management program vs. a brochure [20]), taking part in the intervention was more appealing in one trial (food supplementation vs. nutritional advice [21]), and there was a higher discontinuation probably because of side effects of the drug in one trial (11% in the intervention group and 1% in the control group) [22]. Therefore, this post hoc analysis revealed that differential attrition might have been a serious issue in five trials (6%).

4. Discussion

This study did not find evidence that differential attrition generally occurs in RCTs. We covered a broad spectrum of clinical areas within general and internal medicine, but our results are in line with studies focusing on a single clinical area [7,8]. In five studies, we were able to account for a rather high differential attrition rate based on the way the control group was shaped (e.g., differences in the number of contact moments, additional advantages because of taking part in the intervention). Such potentially anticipatable sources of differential attrition should be taken into account when designing an RCT.

Furthermore, the meta-regression analyses identified “trials in which the control group received an alternative intervention and participants were blinded” as the only factor being related to differential attrition. A relatively higher attrition rate in these control groups could possibly be explained by a lack of perceived treatment efficacy. After this reasoning, however, the attrition rate in control groups with

placebo/sham intervention should be even higher (as perceived treatment efficacy is likely to be even lower), which we did not find. Although interpreting these findings, however, it is important to keep in mind that the estimated average relative attrition rate of all trials combined was essentially indistinguishable from one and had a very narrow CI. This, together with the lack of heterogeneity, suggests that the occurrence of differential attrition in the published literature is mostly a chance finding, unrelated to any particular design factors (e.g., type of control). This is in line with a previous study identifying only a small number of key relationships between patient characteristics and attrition in general [23].

The factors examined in the present study were selected based on their face validity and availability from the published articles, but other possible factors are imaginable (e.g., unintentional differences in the way the groups are treated). It is unlikely, however, that there are unidentified factors related to differential attrition in the present study because the problem of differential attrition was generally quite limited. Another valuable finding of the study at hand, however, is the average overall attrition of 13%. This number can be used as a rule of thumb in sample size calculations of future RCTs within general and internal medicine.

Although the lack of asymmetry in the funnel plot (Fig. 4) does not suggest the presence of publication bias in one specific direction, it could be that there is a publication bias regarding trials in which there is strong differential attrition in either direction. The lack of heterogeneity regarding relative attrition rates is in line with this possibility. This is, however, not problematic and indicates that current review procedures are effective: RCTs in which differential attrition is a problem are less likely to be published (or submitted for publication).

Our conclusions are limited to RCTs published in the 40 general and internal medicine journals included in our sample. A strength of the present study, however, is the systematic way in which trials were randomly sampled from recently published RCTs in these journals and independently scored by two investigators. This is in fact one of the rare cases where a meta-analysis can be said with confidence to provide a representative sample of studies

because there is a clearly defined population of studies from which a random sample of studies was taken. This methodology can be applied to analyze the potential problem of differential attrition in other fields. Although RCTs in this study already covered a broad spectrum of clinical areas, we feel that it might be worthwhile to conduct a comparable study among nonpharmaceutical trials only. It is imaginable, for example, that differential attrition might be a problem in interventions aimed at behavior change in primary prevention (e.g., smoking cessation, increased physical activity). Efficacy trials regarding these types of interventions are normally unblinded. Moreover, participants in these trials are often highly motivated (e.g., they participate despite that they do not experience illness) and, therefore, have high expectations regarding treatment efficacy. Future research is needed to determine whether this leads to a higher degree of differential attrition in these types of trials. In RCTs within general and internal medicine, however, the problem of differential attrition is currently very limited.

References

- [1] Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. *BMJ* 2006;332:969–71.
- [2] Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002;359:781–5.
- [3] Molenberghs G, Kenward MG. *Missing data in clinical studies*. Chichester, UK: Wiley; 2007.
- [4] Stein DJ, Ipser JC, Seedat S. Pharmacotherapy for post traumatic stress disorder (PTSD). *Cochrane Database Syst Rev* 2006;1:CD002795.
- [5] Wortman PM. Differential attrition: another hazard of follow-up research. *Am Psychol* 1978;33:1145–6.
- [6] Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton-Mifflin; 2002.
- [7] Heneghan C, Perera R, Ward AA, Fitzmaurice D, Meats E, Glasziou P. Assessing differential attrition in clinical trials: self-monitoring of oral anticoagulation and type II diabetes. *BMC Med Res Methodol* 2007;7:18.
- [8] Lurie I, Levine SZ. Meta-analysis of dropout rates in SSRIs versus placebo in randomized clinical trials of PTSD. *J Nerv Ment Dis* 2010;198:116–24.
- [9] Rabinowitz J, Levine SZ, Barkai O, Davidov O. Dropout rates in randomized clinical trials of antipsychotics: a meta-analysis comparing first- and second-generation drugs and an examination of the role of trial design features. *Schizophr Bull* 2009;35:775–88.
- [10] Hewitt CE, Kumaravel B, Dumville JC, Torgerson DJ. Assessing the impact of attrition in randomized controlled trials. *J Clin Epidemiol* 2010;63:1264–70.
- [11] King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, et al. Impact of participants and physician intervention preferences on randomized trials. *JAMA* 2005;293:1089–99.
- [12] Puffer S, Torgerson DJ, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medicine journals. *BMJ* 2003;327:785–9.
- [13] Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley; 2008.
- [14] Sterne JAC, Egger M. Regression methods to detect publication and other bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustment*. Chichester, UK: Wiley; 2005: 99–110.
- [15] Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. *Res Synth Meth* 2010;1:112–25.
- [16] R Development Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2010.
- [17] Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.
- [18] Gary TL, Batts-Turner M, Yeh HC, Hill-Briggs F, Bone LR, Wang NY, et al. The effects of a nurse case manager and a community health worker team on diabetic control, emergency department visits, and hospitalizations among urban African Americans with type 2 diabetes mellitus: a randomized controlled trial. *Arch Intern Med* 2009; 169:1788–94.
- [19] Li H, Ma C, Wang G, Zhu X, Peng M, Gu N. Response and remission rates in Chinese patients with bipolar mania treated for 4 weeks with either quetiapine or lithium: a randomized and double-blind study. *Curr Med Res Opin* 2008;24:1–10.
- [20] Chiauzzi E, Pujol L, Wood M, Bond K, Black R, Yiu E, et al. *pain-ACTION-back pain: a self-management website for people with chronic back pain*. *Pain Med* 2010;11:1044–58.
- [21] Martins N, Morris P, Kelly PM. Food incentives to improve completion of tuberculosis treatment: randomised controlled trial in Dili, Timor-Leste. *BMJ* 2009;339:b4248.
- [22] Lower EE, Fleishman S, Cooper A, Zeldis J, Faleck H, Yu Z, et al. Efficacy of dexamethylphenidate for the treatment of fatigue after cancer chemotherapy: a randomized clinical trial. *J Pain Symptom Manage* 2009;38:650–62.
- [23] Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol* 2005;58:13–9.