# The use of older studies in meta-analyses of medical interventions: a survey

Nikolaos A Patsopoulos, John PA Ioannidis

## ABSTRACT

**Background:** Evidence for medical interventions sometimes derives from data that are no longer up to date. These data can influence the outcomes of meta-analyses, yet do not always reflect current clinical practice. We examined the age of the data used in meta-analyses contained within systematic reviews of medical interventions, and investigated whether authors consider the age of these data in their interpretations.

**Methods:** From Issue 4, 2005, of the Cochrane Database of Systematic Reviews we randomly selected 10% of systematic reviews containing at least 1 meta-analysis. From this sample we extracted 1 meta-analysis per primary outcome. We calculated the number of years between the study's publication and 2005 (the year that the systematic review was published), as well as the number of years between the study's publication and the year of the literature search conducted in the study. We assessed whether authors discussed the implications of including less recent data, and, for systematic reviews containing meta-analyses of studies published before 1996, we calculated whether excluding the findings of those studies changed the significance of the outcomes. We repeated these calculations and assessments for 22 systematic reviews containing meta-analyses published in 6 high-impact general medical journals in 2005.

**Results:** For 157 meta-analyses (n = 1149 trials) published in 2005, the median year of the most recent literature search was 2003 (interquartile range [IQR] 2002-04). Two-thirds of these meta-analyses (103/157, 66%) involved no trials published in the preceding 5 years (2001-05). Forty-seven meta-analyses (30%) included no trials published in the preceding 10 years (1996-2005). In another 16 (10%), the statistical significance of the outcomes would have been different had the studies been limited to those published between 1996 and 2005, although in some cases this change in significance would have been due to loss of power. Only 12 (8%) of the meta-analyses discussed the potential implications of including older studies. Among the 22 meta-analyses considered in high-impact general medical journals, 2 included no studies published in the 5 years prior to the reference year (2005), and 18 included at least 1 study published before 1996. Only 4 meta-analyses discussed the implications of including older studies.

**Interpretation:** In most systematic reviews containing meta-analyses of evidence for health care interventions, very recent studies are rare. Researchers who conduct systematic reviews with meta-analyses, and clinicians who read the outcomes of these studies, should be made aware of the potential implications of including less recent data.

**Nikolaos A Patsopoulos** is a Research Associate at the Clinical and Molecular Epidemiology Unit and Clinical Trials and Evidence Based Medicine Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. **John PA Ioannidis** is Professor and Chairman at the Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, with an adjunct appointment at the Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts University School of Medicine, Boston, Mass., USA.

**Correspondence:** Correspondence: John P.A. Ioannidis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece; jioannid@cc.uoi.gr

IN THE SEARCH FOR BEST EVIDENCE, SYSTEMATIC reviewers and clinicians alike often use studies conducted a long time ago. What constitutes "a long time" varies according to scientific topic, the intervention under investigation, methodology, and differences in study population. Whereas the findings of some trials remain valid over time, others become less relevant[1-3] as concomitant therapies and disease management strategies change.[4-6]

To date there has been little discussion of how often older data are used in systematic reviews, or of the reliability and relevance of these data, particularly in regard to systematic reviews containing meta-analyses. To our knowledge, no systematic assessment of the impact of older data on the outcomes of meta-analyses within systematic reviews exists.

In this study we sought to measure how recent is the synthesized evidence in a representative random sample of the Cochrane Database of Systematic Reviews (CDSR), the largest and most comprehensive compilation of systematic reviews on medical interventions.[7-9] We also evaluated whether authors of these studies discussed the age of trials included in their meta-analyses and the implications of their inclusion. Finally, we sought to determine whether meta-analyses published in high-profile journals in 2005 were more likely to use more recent data and if the authors discussed the impact of using less recent data on their findings.

## Methods

**Selecting systematic reviews containing meta-analyses from the CDSR.** We used Intercooled Stata 8.2 (College Station, Tex.) to randomly select 10% of the systematic reviews included in the CDSR, Issue 4 (2005), keeping those that had not been subsequently withdrawn and in which at least 1 meta-analysis had been performed. Selecting only 1 meta-analysis per eligible selected systematic review, we assessed the time since publication of the trials included in the meta-analysis and determined whether the review authors discussed the implications of including less recent trial data. When more than 1 eligible meta-analysis was included in the same systematic review, we retained the meta-analysis that described the primary outcome as defined by the review. If the review included meta-analyses for more than 1 primary outcome and/or more than 1 comparison of interventions, we retained the meta-analysis that encompassed the greatest number of trials. In the case of a tie, we chose the meta-analysis with the greatest cumulative sample size. With further ties for a binary outcome we then selected the meta-analysis with more events. Whenever the same trials had been entered as 2

or more comparisons in the same meta-analysis, we counted them as a single trial and calculated its total sample size. Both authors reviewed the random sample and selected relevant meta-analyses. Discrepancies were resolved by consensus.

**Data extraction and timing of publication.** From each selected meta-analysis, we extracted information on the year of publication of each included trial, the number of participants, and information to calculate the effect size (odds ratio for binary outcomes, standardized or weighted mean difference for continuous outcomes, as specified in each review) and its variance.

Because some authors disseminate their trial results in several articles and meeting abstracts, we consistently used the publication year for the main article, as selected by the Cochrane reviewers. We validated this process by screening the articles and meeting abstracts from a random sample of 50 trials included in our analyses.

For eligible trials in which the Cochrane reviewers had not selected a specific article and year of publication or year for data retrieval, we selected the most recent listed articles or abstracts of the trial; for those studies with entirely unpublished data, we used the year of the last literature search.

We also noted when the last literature search was performed for the systematic review; where this information was unavailable, we recorded the year of last amendment.

**Determining the timing of included studies.** The definition of an "older" study varies across medical specialty and from study to study. For operational purposes, we used the last 5 years (2001-05), 10 years (1996-2005) and 20 years (1986-2005) from the publication year of the CDSR database as pre-specified cutoffs.

**Examining implications of including older trials.** We examined whether the authors of each selected systematic review discussed any implications relating to the fact that some or all of the trials included in their meta-analyses were older, and whether systematic reviews discussing these implications had included, overall, older trials than those that did not discuss any such issues.

**Sensitivity analyses.** We conducted sensitivity analyses addressing the 5-, 10- and 20-year estimates using the year of the literature search rather than the CDSR issue date (2005) as a reference year. Using the latter reflects how recent the evidence is for clinical practice at the time of the CDSR issue publication, while using the year of the literature search subtracts the time lag between the literature review and its publication in 2005.

We also assessed the impact on meta-analyses of including only studies published in the last decade (1996-2005). Specifically, we recorded how many meta-analyses included such studies, and how many would have reached different inferences for the statistical significance of the summary effect ($p < 0.05$ or $p \geq 0.05$) had only these studies been included.

**Examining systematic reviews published in major medical journals.** We also assessed systematic reviews containing meta-analyses published in 6 major medical journals (New England Journal of Medicine, Journal of the American Medical Association, The Lancet, British Medical Journal, Annals of Internal Medicine and PLoS Medicine) between July and December 2005. The examined journals publish systematic reviews of broad medical interest, thus reducing the risk of evaluating topics on few specialties. We followed a protocol similar to that described above for the selection of meta-analyses from systematic reviews in the CDSR.

The year (late 2005) was chosen to be the same as the issue of CDSR that we evaluated. We included only systematic reviews containing at least 1 meta-analysis. When more than 1 meta-analysis of interventions and/or outcomes existed, we followed the same procedure as for the CDSR systematic reviews to select 1 meta-analysis per article. From each eligible meta-analysis we extracted the year of the literature search, the number of synthesized trials, and the number of trials published in the last 5 years (2001-05), 10 years (1996-2005) and 20 years (1986-2005). We did not collect information on sample size per trial and effect size because these were not sufficiently standardized to allow consistent analyses. Otherwise, we evaluated the proportion of recent data and whether the implications of including older data were discussed by review authors, as described for the analysis of systematic reviews from the CDSR.

**Statistical analysis.** Summary results were calculated using the DerSimonian and Laird random effects model[10] allowing for variability across trials. Comparison of the publication years of the synthesized trials years was made with Mann-Whitney U test. P values are two-tailed. Analyses were conducted with Intercooled Stata 8.2 (College Station, Tex.).

## Results

**Systematic reviews and meta-analyses selected from the CDSR.** We randomly sampled 165 of 1651 eligible systematic reviews from the CDSR (Fig. 1). Of these, 8 were excluded (6 had been withdrawn, and 2 contained no quantitative synthesis). Of the remaining 157 system-

atic reviews, the selected outcomes were binary in 133 meta-analyses and continuous in 24 meta-analyses. (See Online Appendix 1 for a complete list of the CDSR systematic reviews and meta-analyses included in this study.)

Overall, 1149 trials (1 650 701 participants) were included in the meta-analyses for the primary outcomes for included systematic reviews. There was a median of 5 trials included in each meta-analysis (IQR 3-7) and of 617 (IQR 227-1711) participants.

The median year (IQR) of the last literature search across the 157 systematic reviews was 2003 (2002-04). There were 19 reviews whose date of last search was in 2005, and 43 whose last search was in 2004. In another 11 reviews the last year of literature search was before 2000, and in 6 of these the most recent included trial had been published at least 3 years earlier. The representation of recent data increased slightly when the reference was shifted to the year of the last literature search rather than 2005 (i.e., the year of publication of the systematic review).

Of the random sample of 50 trials (of the total 1149 trials encompassed by the 157 meta-analyses included in our review) 44 had the same publication year; for 1 trial, the publication years spanned 2 consecutive years, for another, publication years spanned 3 consecutive years, and in 4 trials publication years spanned a longer period.

Overall, few meta-analyses consisted of data from trials published in the 5 years prior to the publication in the CDSR (2005), and only a quarter of the trials they synthesized had been published within 5 years of the last literature search (Table 1). One-third of meta-analyses included data from trials that had been published in the last 10 years, and two-thirds had been published within 10 years of the literature search. A 20-year window usually captured most or all of the trials that contributed data to the included meta-analyses.

Almost one-third (30%, 47/157) of meta-analyses included no trials published in the period 1996-2005, while only 2 meta-analyses (1%, 2/157) consisted entirely of data from trials published in the same period (Table 2). The respective proportions were 18% (29/157) and 55% (87/157) for a 10-year time frame from the year of the systematic reviews' last literature search.

Of the 157 meta-analyses surveyed, 47 included only trials published before 1996. Of the remaining 110 meta-analyses, 21 included trials that were all published in the last decade. Among the 89 meta-analyses that included trials published in the last decade and earlier, exclusion of the data published before 1996 changed the level of statistical significance of the summary effects in 16; in 9 cases the summary effects became non-statistically significant, and the opposite change was seen in 7 cases. The summary odds ratio for these 16 meta-analyses changed by a median of 23% (IQR 9%-41%).

**Discussion of implications of including less recent studies.** Only 12 of 157 systematic reviews (8%) discussed the possible implications of including older trials. These included meta-analyses related to pregnancy and neonatal medicine (4), neuropsychiatric disease (4), pulmonary disease (2), diabetes mellitus (1) and chemotherapy for bladder cancer (1). The authors of 2 of these reviews concluded that the year of publication was unlikely to matter; 10 reviews expressed some concern, and 2 clearly stated the importance of revisiting the clinical question with new trials. Trials included in these 12 meta-analyses were significantly older than those included in the other 145 meta-analyses that did not discuss the age of the trials ($p = 0.003$).

Most of the systematic reviews that included trials published before 1986 (61/70) and 124 of the 136 systematic reviews that included trials published before 1996 did not discuss the age of the included trials.

**Systematic reviews published in major medical journals.** Twenty-five systematic reviews containing meta-analyses were published in 2005 in the 6 high-impact general medical journals, of which 22 had data available on the publication year of the synthesized trials (Online Appendix 2). The median (IQR) of the literature search year, reported in 15 of the 22 reviews, and the number of synthesized studies were 2005 (2004-05) and 11 (8-15), respectively. Median proportions (IQR) of trials published in the last 5, 10, or 20 years since 2005 were 39 (17-50), 68 (45-90), and 100 (92-100), respectively. Two meta-analyses had no trials published in the previous 5 years (2001-05), while all had at least 1 trial published in the last decade (1996-2005). Eighteen meta-analyses included at least 1 study published before 1996.

Four systematic reviews discussed the implications of including less recent data in their meta-analyses. One of these also performed a sensitivity analysis that included only trials published in the last decade. The median year of publication for the trials included in these 4 systematic reviews showed no statistically significant difference than in the other 19, which did not discuss the age of the trials (p = 0.46). Overall, 7 of the 9 systematic reviews that included trials published before 1986 did not discuss the implications of including less recent trials.

## Discussion

Our analysis suggests that, even with the best intentions, evidence-based medicine has to rely on less recent evidence. Even when results were corrected for the year of the last literature search, few systematic reviews containing meta-analyses in the CDSR included trials published in the preceding 5 years. Almost a third of these included no trials published in the last decade, and in another 10% the statistical significance of the result related to the study primary outcome would have been different had the data been limited to the last decade. Most of the systematic reviews did not address the implications of including less recent data.

Meta-analyses published in high-profile peer-reviewed journals tend to address newer interventions than the average CDSR review. Accordingly, almost all of them included some trials published in the last 5 years, and all of them included some trials published in the last decade. Nevertheless, even in these meta-analyses, the large majority also included 1 or more older trials, and very few discussed the implications of including older evidence.

Figure 1. Flow diagram of selection of systemic reviews and meta-analyses
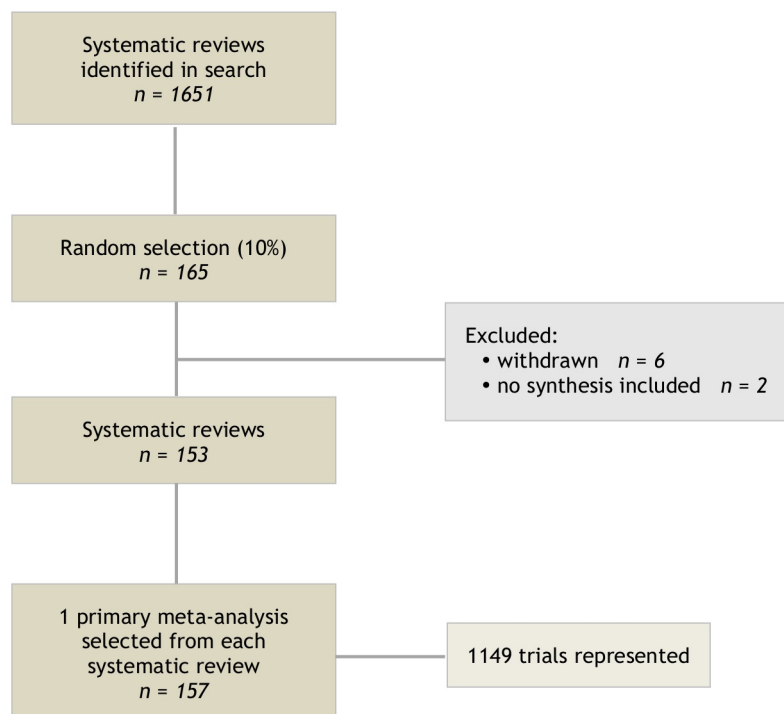
**Table 1: Absolute number or proportion of trials included in meta-analyses in which year of publication was within 5, 10 or 20 years from the year of systematic review publication in the CDSR (2005) or the year of last literature search**

|  | Last 5 years | Last 10 years | Last 20 years |
|---|---|---|---|
| **Median year of publication in the CDSR (2005)** | | | |
| Absolute number of trials (IQR) | 0 (0–1) | 2 (0–4) | 5 (2–8) |
| Proportion of trials (IQR) | 0 (0–17) | 33 (0–71) | 100 (60–100) |
| **Median year of last literature search** | | | |
| Absolute number of trials (IQR) | 1 (0–4) | 3 (2–6) | 5 (3–9) |
| Proportion of trials (IQR) | 25 (0–50) | 69 (33–100) | 100 (92–100) |

CDSR = Cochrane Database of Systematic Reviews; IQR = interquartile range

**Table 2: Absolute numbers and proportion of included meta-analyses with either no trial or all trials published in the las 5, 10, or 20 years from the reference year (n = 157)**

|  | Reference year* | Last 5 years | Last 10 years | Last 20 year |
|---|---|---|---|---|
| **No trials published in this time frame** | 2005 | 103 (66%) | 47 (30%) | 8 (5%) |
|  | Last literature search | 71 (45%) | 29 (18%) | 6 (4%) |
| **All trials published in this time frame** | 2005 | 2 (1%) | 2 (1%) | 21 (13%) |
|  | Last literature search | 39 (25%) | 87 (55%) | 104 (66%) |

* The year used to define the time intervals. 2005 is the publication year of the CDSR under use, and "last literature search" refers to the year each review's aut searched for new studies.

Typically, the lack of recent evidence did not result from the CDSR systematic reviews being out of date; in fact, the majority of systematic reviews that we analyzed had been updated in the last 2 years. Nonetheless, few systematic reviews discussed the implications of the time of publication on the relevance of the evidence.

Evidence should not be undervalued simply because of its age. The amount of data, regardless of year of publication, is limited for most health care topics,[11-13] and we do not have the luxury of discarding trials simply because of their calendar year. In the case of topics for which well-designed old clinical trials are still relevant and conclusive, it is imprudent, and even unethical, to conduct new trials.

Occasionally, earlier published results may differ from those reported in later publications.[14-16] This may reflect bias,[17-19] time-dependent efficacy (e.g., when the treatment benefit decreases with longer follow-up),[20] quality differences,[21,22] or chance. For example, in the case of vitamin E supplementation for prevention of morbidity and mortality in preterm infants, the review authors suggest caution in interpreting and applying current evidence. The available data span over a decade (1991-2002), a period in which many advances were made in the field of preterm care.[23] However, one cannot generalize. Less recent trials are not necessarily of worse quality,[24,25] smaller,[26] or less externally valid[4,5] than newer ones.[2] Each topic needs a careful case-by-case scrutiny of whether the available evidence is relev-

ant to current practice. The availability of evidence is sometimes further restricted by the lack of standardized outcomes across trials. Selective reporting of "positive" outcome results is an added threat.[27-29]

Some limitations should be discussed. Although we used a standardized approach to select the year of publication, a trial may be in progress for many years before any results are published. Most trials do not specify when they started and completed enrolment and follow-up. Efficacy trials may take 3 to 10 or more years from the start of enrolment to publication.[17] Therefore, the proportion of recently conducted trials is likely even smaller than what we report on the basis of publication year.

Second, we used the CDSR for our primary analyses because it is widely considered the most all-encompassing and up-to-date source for current evidence on health care interventions. However, even the CDSR represents work in progress, and it does not capture all interventions.[30] Furthermore, some review authors may choose to exclude, a priori, less recent studies, especially in fast-moving areas of research, by restricting search years or requiring the reporting of methodological quality characteristics.

Our evaluation of systematic reviews published in medical journals was unavoidably more restricted, since some information (such as primary outcome) is not standardized and is readily available in the same detail as in the CDSR systematic reviews.

We should also caution that decision-making based on nominal statistical significance is precarious.[31,32] A change in statistical significance does not mean that the estimated effect size is altered beyond chance. Neither can it be attributed with certainty to less recent studies, since many underlying factors (including chance alone) contribute to the uncertainty of the effect estimate. Given that most meta-analyses had very limited data overall, there was large uncertainty in the estimated effect size in recent compared with less recent published data. Direct comparisons of recent against less recent data would be underpowered to show even major differences in effect sizes in these meta-analyses.

However, some empirical evidence suggests that, in some fields, smaller treatment effects may be encountered in more recent trials than in earlier research.[1,14-16,33] In the present evaluation, in the meta-analyses in which the formal statistical significance status of the summary effect changed with the exclusion of less recent data, the median change in the odds ratio was also 23%. This is a considerable change, given that most medical interventions have modest effects.

Acknowledging these caveats, our survey suggests that even though the CDSR reviews are frequently updated, evidence from very recently published studies for most health care interventions is scant. Although less recent studies should not be discarded, clinicians should interpret medical evidence with attention to the applicability and relevance of these studies to current clinical practice. If evidence on a specific topic is considered to be outdated or missing, and the review question remains salient, the scientific community should be sensitized toward conducting relevant targeted studies.

## References

1. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA.* 2005;294(2):218–228.

2. Poynard T, Munteanu M, Ratziu V, Benhamou Y, Di Martino V, Taieb J, et al. Truth survival in clinical research: an evidence-based requiem? *Ann Intern Med.* 2002;136(12):888–895.

3. Hall JC, Platell C. Half-life of truth in surgical literature. *Lancet.* 1997;350(9093):1752.

4. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials.* 2006;1(1):e9.

5. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply"? *Lancet.* 2005;365(9453):82–93.

6. Ioannidis JPA. Indirect comparisons: the mesh and mess of clinical trials. *Lancet.* 2006;368(9546):1470–1472.

7. Bero L, Rennie D. The Cochrane Collaboration. Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA.* 1995;274(24):1935–1938.

8. Clarke M, Langhorne P. Revisiting the Cochrane Collaboration. Meeting the challenge of Archie Cochrane–and facing up to some new ones. *BMJ.* 2001;323(7317):821.

9. Grimshaw J. So what has the Cochrane Collaboration ever done for us? A report card on the first 10 years. *CMAJ.* 2004;171(7):747–749.

10. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177–188.

11. Muir Gray JA. *Evidence-based healthcare: how to make health policy and management decisions.* London: Churchill Livingstone: 1997.

12. Djulbegovic B, Loughran TP, Hornung CA, Kloecker G, Efthimiadis EN, Hadley TJ, et al. The quality of medical evidence in hematology-oncology. *Am J Med.* 1999;106(2):198–205.

13. Mallett S, Clarke M. The typical Cochrane review. How many trials? How many participants. *Int J Technol Assess Health Care.* 2002;18(4):820–823.

14. Ioannidis J, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci U S A.* 2001;98(3):831–836.

15. Trikalinos TA, Churchill R, Ferri M, Leucht S, Tuunainen A, Wahlbeck K, et al. EU-PSI project. Effect sizes in cumulative meta-analyses of mental health randomized trials evolved over time. *J Clin Epidemiol.* 2004;57(11):1124–1130.

16. Gehr BT, Weiss C, Porzsolt F. The fading of reported effectiveness. A meta-analysis of randomised controlled trials. *BMC Med Res Methodol.* 2006;6:25–26.

17. Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA.* 1998;279(4):281–286.

18. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ.* 1997;315(7109):640–645.

19. Hopewell S, Clarke MJ, Stewart L, Tierney J. Time to publication for results of clinical trials. *Cochrane Database of Systematic Reviews.* . 2007;(Issue 2)

20. Ioannidis JP, Cappelleri JC, Sacks HS, Lau J. The relationship between study design, results, and reporting of randomized clinical trials of HIV infection. *Control Clin Trials.* 1997;18(5):431–444.

21. Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA.* 1998;279(14):1089–1093.

22. Schulz KF. Subverting randomization in controlled trials. *JAMA.* 1995;274(18):1456–1458.

23. Brion LP, Bell EF, Raghuveer TS. Vitamin E supplementation for prevention of morbidity and mortality in preterm infants. *Cochrane Database of Systematic Reviews.* . 2003;(Issue 4)

24. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med.* 2001;134(8):663–694.

25. Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med.* 2004;141(10):781–788.

26. Chan A, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet.* 2005;365(9465):1159–1162.

27. Chan A, Krleza-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ.* 2004;171(7):735–740.

28. Chan A, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA.* 2004;291(20):2457–2465.

29. Chan A, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ.* 2005;330(7494):753.

30. Mallett S, Clarke M. How many Cochrane reviews are needed to cover existing evidence on the effects of health care interventions. *ACP J Club.* 2003;139(1):A11.

31. Ioannidis JPA. Why most published research findings are false. *PLoS Med.* 2005;2(8):e124.

32. Djulbegovic B, Hozo I. When should potentially false research findings be considered acceptable? *PLoS Med.* 2007;4(2):e26.

33. Hauben M, Reich L, Van Puijenbroek EP, Gerrits CM, Patadia VK. Data mining in pharmacovigilance: lessons from phantom ships. *Eur J Clin Pharmacol.* 2006;62(11):967–970.