A Simple Method For Detecting And Adjusting Meta-Analyses For Publication Bias

F Richy, J Reginster

Citation

F Richy, J Reginster. *A Simple Method For Detecting And Adjusting Meta-Analyses For Publication Bias*. The Internet Journal of Epidemiology. 2005 Volume 3 Number 2.

Abstract

Background: Publication bias is widely recognized in the field of evidence-based medicine as a threat on the validity of the results provided by meta-analyses. Several methods, of various complexities, have been developed, but their systematic implementation is limited by the criticisms formulated against some of their assumptions. We developed an innovative, easy to implement, method able to simultaneously detect and correct for publication bias in studies using dichotomised outcomes (relative risk, Odds ratio).

Methods: The method is based on the sum of the moments of forces (defined by $[\text{Ln}(RR_i - \text{Ln}(RR_{global})] \times [\text{var}(\text{Ln}(RR_i))]^{-1}$)

weighting on the global estimate provided by classic meta-analysis. When a significant imbalance is found (i.e. when the required moment of forced exceeds the 95% confidence interval around equilibrium), the system can be counterbalanced by generating the “missing studies”, taken from the mirrored set of actual studies. The method was compared to the linear regression approach using a simulated dataset in which we intentionally induced publication bias, and afterwards on two datasets from previous meta-analyses by our group.

Results: Our method provided trends for a more sensitive detection of publication bias when compared to the linear regression method in the simulated dataset. Our method also seemed to be more powerful to ascertain that no publication bias was present in actual data (e.g. its statistical power was higher).

Implications: Despite these encouraging observations, our method remains to be validated in different datasets and against other approaches. If it appears to be effective and reliable, it might be systematically applied after excluding confounding outcomes by preliminary meta-regression.

BACKGROUND

The use of a biased set of studies in terms of completeness or effect magnitude has been regularly cited as the main criticism against meta-analysis, since its development in the late seventies[1,2,3,4]. Indeed, using a non-representative set of studies is prone to lead to and overestimation of the “true” intervention effect. The initial report of publication bias dates to 1959 when Sterling noticed that 97% of articles published in four journals reported statistically significant findings, raising the likelihood that studies lacking significant differences were not being published[5]. Funnel plot techniques for evaluating the probability of publication bias have initiated in 1984 by Light and colleagues[6] and have been introduced for the first time in formal research in 1988 by Vandenbroucke and colleagues[7]. This graphic approach to publication bias has been developed, mainly thanks to the work of Matthias Egger and colleagues[3,8]. Its grounds are scatter plots of the treatment effects estimated from individual studies (horizontal axis) against some measure of study size (vertical axis). Because only precision in estimating the underlying treatment effect increases as a study's sample size increases, effect estimates from small studies scatter more widely at the bottom of the graph, with the spread narrowing among larger studies. In the absence of bias the plot therefore resembles a symmetrical inverted funnel. Asymmetrical funnel plots may indicate publication bias or be due to exaggeration of treatment effects in small studies of low quality (small study effects). Indeed, we have arguments favouring an additional probabilistic nature of

small study effect, distinct from publication bias[9].

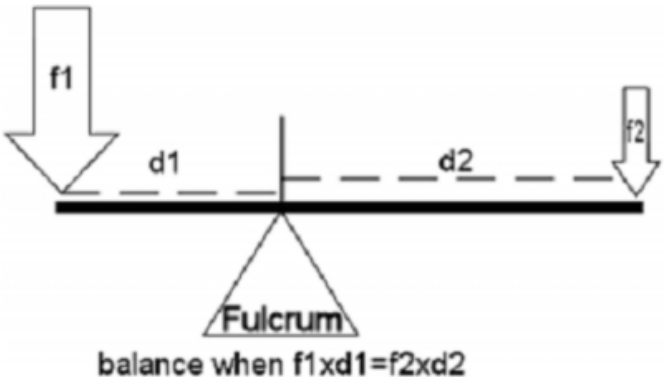
Sterne and colleagues[3,10] and Thornton and colleagues[2] have recently reviewed the available methods to detect and, further, to adjust the global estimates of meta-analyses, for publication bias. Several researchers have proposed methods, of various complexities and requirements, to this end. The association between treatment effect size and its standard error, unrelated to sample-size, is the keystone of two statistical methods, a rank correlation test[11] and a regression method[8], which have been proposed as a means of avoiding the subjectivity associated with visual assessment of funnel plots. However the validity of these methods has been questioned[12,13,14]. Other methods, like trim and fill[15] have been proposed to adjust the global estimate for funnel plot asymmetry. Trim and Fill builds on the key idea behind the funnel plot; that in the absence of bias the plot would be symmetric about the summary effect. If there are more small studies on the right than on the left, the concern is that studies may be missing from the left. The Trim and Fill procedure imputes these missing studies, adds them to the analysis, and then re-computes the summary effect size. However, simulation studies have found that the trim and fill method may suffer from low specificity, e.g. it might detect "missing" studies in a substantial proportion of meta-analyses, even in the absence of bias[16]. Thus there is a danger of over correcting nonexistent bias in response to funnel plot asymmetry arising from nothing more than random variation. Thus, the current consensus on correction of treatment effect estimates for bias reflects that it should be avoided as such corrections may hardly depend on the assumptions made[10,17].

The aim of this paper was to describe an innovative, easy and systematically applicable method for simultaneously detecting and correcting for publication bias in meta-analyses of published data.

METHODS

Our method uses the principle of "moment of force" (MF) that we have all seen in physics classes. The moment of a force is the multiplication of the force size by its distance to the fulcrum of the considered lever. The moment of the f1 force (in Newtons) is f1xd1 Newtons-metre. The balance is reached when the moment of the f_1 (MF1) force equals the moment of the f_2 (MF2) force: $\sum[(f1 \times d1) + (f2 \times d2)] = 0$ (figure 1).

Figure 1



Our approach for detecting and correcting for publication bias is based on this simple principle. In a classic funnel plot graph of relative risks or odds ratios, the Ln(estimate) is plotted against its precision (1/Std ERR(Ln estimate)). In our model, the fulcrum is located below the value of the Ln(global estimate) provided by the meta-analysis. Each study has a specific moment of force on the system, represented by the multiplication of its precision (the force applied to the system) f_i by its distance to the global estimate (d_i-d□). The main assumption in our model is that, in absence of publication bias, the sum of all the moments of force is null, and the system is in equilibrium: $X = -\sum_{i=1}^{i=k}[f_i \times (d_i - d□)] = 0$. When the sum of the moments of force is unequal, we can easily obtain the "missing" one required to balance the system. The core concept of our formal analysis is the following. If the obtained "missing" moment of force overtakes the 95% non-parametric confidence interval of the distribution of all MFs, we can then postulate that a significant publication bias is present, whatever the symmetry of the funnel plot. In such a case, the global estimate of the meta-analysis can be easily corrected by adding one or more "negative" studies counterbalancing the system and re-performing the meta-analysis including this additional data.

A simple approach to determine the number of missing studies is to divide the parameter X by highest moment of force among existing "negative" studies: . n = Abs(X / min(MFi). The funnel plot can be filled using the distribution of the mirrored positive studies. For n studies, the percentiles of the couples di-precision to be considered are multiples of: k = (100 / n). For instance, for n=three negative studies missing, the parameter k will be 100/6=16.67. The considered percentiles of the mirrored positive d_i's would be 100, 66.67 and 33.33 and the percentiles of the corresponding □$_i$ would be 0, 33.33 and

66.66. To achieve an exact equilibrium of the system, the smallest d_i can be rounded.

We compared our results with a formal statistic test of funnel plot asymmetry[8], shown to be more powerful than the rank correlation method[3]. In this linear regression method, the standard normal deviates of the Ln(RRi) (defined by Ln(RRi) divided by its standard error) are regressed against their precisions (1/Std error of the Ln(RRi). The main assumption is that, in absence of publication bias, the linear regression scatters about a line that runs through the origin at SND=0, at p<0.10.

The validation of a method to detect publication bias remains difficult in absence of a "gold standard" set of data, in which we can formally exclude a prevalent publication bias. An approach would consist in working on the basis of couples of data including pairs of large RCTs and their corresponding meta-analyses. Notwithstanding, this approach suffers some limitations. Using such pairs induces a variability in the appraisal of their concordance since large RCT may themselves be victims of publication bias, as outlined with the recent hype on paroxetine [18], and because of the lack of association between precision and sample-size[8]. Furthermore, small studies and large ones may not reflect the same treatment effect, due to risk differences among included patients, comedications, duration, etc[3].

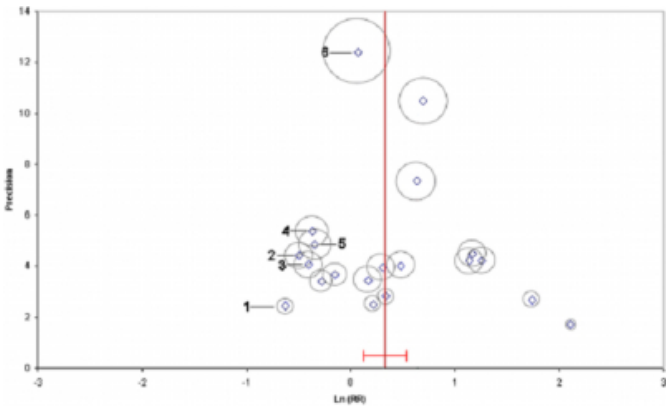
In order to test our method in a evidence-based way, we have initially applied it to a set of 20 studies that we have generated using an iterative algorithm simulating the estimates and the precisions of twenty trials, under a condition of strict equivalence (difference <0.0001) of the moments of force below and above the corresponding global estimate given by meta-analysis (figure 1). As a second step, we applied our method to actual datasets extracted from our previous works on glucosamine and chondroïtin[19], and D-analogs[20].

All development operations were done using Microsoft Excel and Visual Basic (® Microsoft Corporation 1985-2005) and Statistica 7.0 (Statsoft, France, 2005). The meta-analytic calculations were performed using Comprehensive meta-analysis 1.0.25 and 2.0(Biostat, USA, 2002-2006).

RESULTS

SIMULATED SET OF DATA

Figure 2



The corresponding meta-analysis statistics provided:

Figure 3

RR (95% CI)	Ln (RR) (95% CI)	Test for Association	Test for Heterogeneity (Hedges & Olkin)	Test for Publication bias
1,40 (1.14; 1.71)	0.34 (0.13; 0.54)	Z=3,24 p=0,003	Q=37,5 Df=19 p=0,06	X=0,000005 95%CI= (-3.81; 3.90) Intercept= 0.62 (p=0.65, NS)

We have sequentially removed the studies #1 to #6 at the extreme left side of the distribution to simulate an increasing publication bias favouring the treatment effect. We have iteratively computed the corresponding global estimate, tests for association and heterogeneity, and compared the linear regression test for publication bias to our method.

Figure 4

	RR (95% CI)	Ln (RR) (95% CI)	Test for Association	Test for Heterogeneity (Hedges & Olkin)	Test for Publication bias
One study removed (1)	1,44 (1.17;1.76)	0.36 (0.16; 0.57)	Z=3.44 p=0.001	Q=35,48 Df=18 p=0,008	X=-2.36 (NS) 95%CI= (-3.80; 3.90) $n_{missing}$=0,62
					Intercept: -0.46 (p=0.87, NS)
Two studies removed (1-2)	1.5 (1.21;1.85)	0.4 (0.19; 0.62)	Z=3.76 p=0.0003	Q=32.4 Df=17 p=0.013	X=-6.02 (S) 95% CI= (-3,80; 3.90) $n_{missing}$ =1,58
					Intercept: -0.22 (p=0.93, NS)
Three studies removed (1-3)	1.58 (1.27;1.96)	0.45 (0.24; 0.67)	Z=4.11 p<0.0001	Q=29.7 Df=16 p=0.02	X=-9.83 (S) 95%CI= (-3,31; 3,90) $n_{missing}$ =2,96
					Intercept: -0.34 (p=0.89, NS)
Four studies removed (1-4)	1.66 (1.33;2.08)	0.51 (0,28; 0,73)	Z=4.43 P<0.0001	Q=27.49 Df=15 p=0.025	X=-13,1 (S) 95% CI= (-3,26; 3,90) $n_{missing}$ =4,03
					Intercept: -0,27 (p=0.92, NS)
Five studies removed (1-5)	1.73 (1.37;2.17)	0.54 (0.31; 0.78)	Z=4.65 P<0.0001	Q=26.21 Df=14 p=0.024	X=-15.23 (S) 95% CI= (-3.26; 3.90) $n_{missing}$ =4.67
					Intercept: 0,27 (p=0.91, NS)
Six studies removed (1-6)	1.91 (1.48;2.45)	0.65 (0.39; 0.90)	Z=4.99 P<0.0001	Q=25.35 Df=13 P=0.02	X=-18,50 (S) 95% CI= (-3,02; 3,90) $n_{missing}$ =6,12
					Intercept: -2,86 (p=0.34, NS)

After removal two studies and more, and using our method, the system appeared to be no more balanced, since a moment of force of -6,02 was needed to adjust for equilibrium, which was out of the 95% confidence interval around the balance point (-3,80; 3,90). Using this dataset, our method displayed a better predictive ability compared to the linear regression method, which was not able to detect any publication bias.

When two negative studies were actually removed from the original set of 20 studies, the number of missing studies provided by our method was 1.58, rounded to 2. Accordingly, the following data were added:

Figure 5

	di	wi	MF	Adjusting for exact balance
Missing study #1	Percentile 100 of the positive di's =-1,77	Percentile 0 of the corresponding wi's =1,71	-3,01	MF=3,01-0,55=2,66 Di=-1,55 Wi=1,71
Missing study #2	Percentile 50 of the positive di's =-0,80	Percentile 50 of the corresponding wi's =4,21	-3,36	
Moment of force added			-6,37	-6,02

The corresponding funnel plot became:

Figure 6

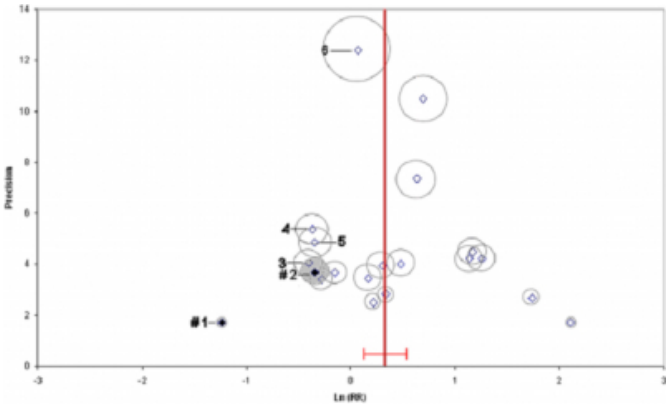


Figure 7

RR (95% CI)	Ln (RR) (95% CI)	Test for Association	Test for Heterogeneity (Hedges & Olkin)	Test for Publication bias
1,41 (1.15; 1.72)	0.34 (0.14; 0.54)	Z=3,26 p=0,002	Q=35,4 Df=19 p=0,01	X=0,008 95%CI= (-3,80; 3,90)

In the situation of a "massive" publication bias (six studies removed), the following studies were added:

Figure 8

	di	Wi	MF	Adjusting for exact balance
Missing study #1	-1,77	1,71	-3,03	MF=-3.03-(18.5-15,62-=-5,91 Di=-3,45 Wi=1,71
Missing study #2	-1,24	2,74	-3,39	
Missing study #3	-0,86	3,63	-3,12	
Missing study #4	-0,80	4,21	-3,36	
Missing study #5	-0,34	4,32	-1,46	
Missing study #6	-0,19	6,41	-1,24	
Moment of force added			-15,62	-18.5

The corresponding funnel plot became:

Figure 9

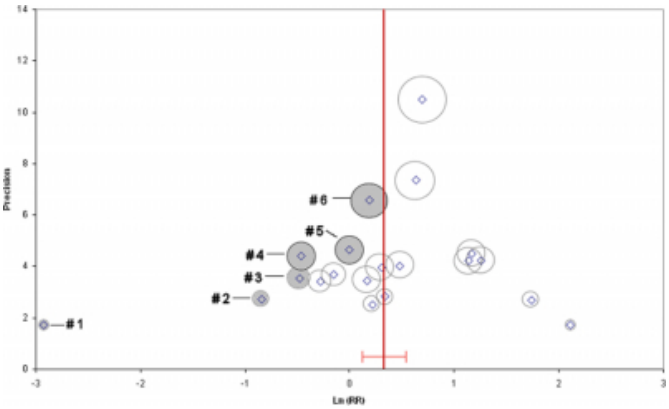


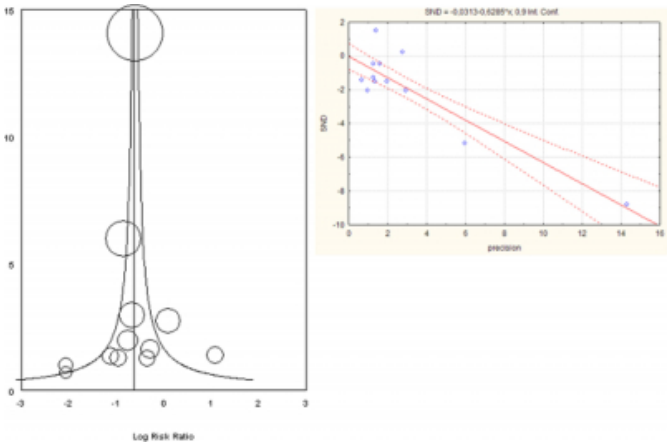
Figure 10

RR (95% CI)	Ln (RR) (95% CI)	Association test	Heterogeneity test (Hedges & Olkin)	Publication bias
1,40 (1.13; 1.74)	0.34 (0.12; 0.55)	Z=3,06 p=0,003	Q=56.06 Df=19 P<0.0001	X=0,005 95%CI= (-5,9; 3,90)

APPLICATION OF THE METHOD TO EXISTING DATASETS

The first set of data we used was the one from our meta-analysis on glucosamine and chondroïtin[19]. Using this dataset and the outcome "responders to treatment vs placebo", the use of the linear regression method provided borderline significant results: intercept=1.67 at p=0.107.

Figure 12



Our method provided a general moment of force of X=-5.59. The 95% confidence interval around the equilibrium was (-9.9; 5.55), reflecting a non-significant publication bias. The number of missing studies was 5.59/9.90=0.56.

When applying the two methods to our dataset on the antifracture efficacy of D-analogs alfacalcidol and calcitriol[20]:

{image:12}

The linear regression method led to non-significant results: intercept=-0.03 (0.94) and our method provided a global moment of force of 5.78; the 95% confidence interval around the equilibrium was (-7.56; 7.70), also reflecting a non-significant publication bias. The number of estimated missing studies was: 5.78/7.70=0.75.

DISCUSSION

The concerns about publication bias have been regularly disputed since the mid-eighties. As this bias occurs in a place where it may bias the two main components of evidence-based medicine (the RCTs and their meta-analyses), it has been argued that all results of meta-analysis with asymmetrical funnel plots should be regarded with caution[21,22]. Although the development of methods for assessing publication bias has a reasonably long history, these methods are rarely used in practice. One possible reason for the lack of uptake of methods to deal with publication bias is that previous approaches have involved modelling methods that are difficult to implement and require lengthy calculations.

Since the remarkable publications by Matthias Egger and colleagues in the Journal[8,23,24,25,26], the potentially serious consequences of such publication bias have been realised for some time, and there have been repeated calls for worldwide registration of any ongoing clinical trial[27,28,29,30]. Unfortunately, it is unlikely that this will be widely instituted in the foreseeable future. This is why publication bias analysis and correction methods have to become essential, systematic parts of meta-analyses.

Some cautionary remarks are needed in assessing tests for publication bias. Since the vast majority of methods is based on the lack of symmetry in the funnel plot, and asymmetry might be due to factors other than publication bias[3,31], the results produced by some methods may not always reflect detection of publication bias nor correction for it.

In the present paper, we developed and applied an innovative, simple method able to detect, and correct for publication bias that does not rely on symmetry or correlation assumptions. On the basis of the present evidence, we were able to find that it might have several advantages compared to the linear regression or other much more complicated methods. It may be the simplest, but highly effective, approach. Indeed, much more sophisticated methods have been developed, including trim-and-fill[15] and rank correlation methods[11] and other likelihood maximisation methods which may be beyond the expertise of non-mathematicians involved in meta-analysis. Importantly, our approach does not postulate that the distribution of the funnel plot has to be strictly symmetrical. We have to keep in mind that when analysing funnel plots, we are generally dealing with restricted numbers of studies, assessing a true treatment effect with a random error around it. The probability of getting a perfectly symmetrical funnel plot is actually low, even in absence of publication bias, and this fact limits the power of tests based on this assumption. Our model bypasses this limitation, the only hypothesis

being that the relative weights of positive and negative evidence are equal, whatever their distribution and level of significance. Using a non-parametric construction of the confidence interval around the fulcrum may be more adapted to variable distribution patterns compared to formally compare the obtained values to a predefined distribution (e.g. chi-square). Last but not least, it seems to be more powerful in predicting publication bias than the regression method while it allows correcting the meta-analysis estimates easily. However, the nature of the selection mechanism, the range of variances of the effect size estimates, and the true underlying effect size are all observed to be influential in determining the power of the test. in many of the configurations in which there was low power, there was also relatively little bias in the summary effect size estimate. Nonetheless, the authors stated that the test should be interpreted with caution in small meta-analyses. In particular, bias cannot be ruled out if the test is not significant.

Any adjustment method should be used primarily as a form of sensitivity analysis together with metaregression techniques. A correction of the global estimate should be done only when covariants appear to be non-significantly impacting it, since some may be confounding factors.

Finally, it should be noted that articles on publication bias are susceptible to such discrimination themselves. That is, papers seeking to report the failure to find publication bias may face a lesser chance of being submitted and accepted for publication.

CORRESPONDENCE TO

Professor Florent RICHY University of Liège CHU, B23 B-4000 SART-TILMAN BELGIUM Tel: +32 4 366 2581 Fax: +32 4 366 2812 Email: florent.richy@ulg.ac.be

References

1. Thompson SG, Pocock SJ. Can meta-analyses be trusted? Lancet 1991; 338: 1127-30.
2. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 2000; 53: 207-16.
3. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000; 53: 1119-29.
4. Begg CB. A comparison of methods to detect publication bias in meta-analysis by P. Macaskill, S. D. Walter and L. Irwig, Statistics in Medicine, 2001; 20:641-654. Stat Med 2002; 21: 1803; author reply 4.
5. Sterling T. Publication decisions and their possible effects on inferences drawn from tests of significances. Am Stat Assoc 1959; 54: 30-4.
6. Light R, Pillemer D. Summing Up: the Science of Reviewing Research. Cambridge, 1984.
7. Vandenbroucke JP. Passive smoking and lung cancer: a publication bias? Br Med J (Clin Res Ed) 1988; 296: 391-2.
8. Egger M, Davey Smith G, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. Bmj 1997; 315: 629-34.
9. Richy F, Ethgen O, Bruyere O et al. From sample size to effect-size: Small study effect investigation (SSEi). The Internet Journal of Epidemiology 2004; 1.
10. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. Bmj 2001; 323: 101-5.
11. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994; 50: 1088-101.
12. Naylor CD, Tu JV. Progress in reducing inpatient mortality from acute myocardial infarction is slow. Bmj 1997; 315: 1462.
13. Irwig L, Macaskill P, Berry G et al. Bias in meta-analysis detected by a simple, graphical test. Graphical test is itself biased. Bmj 1998; 316: 470; author reply -1.
14. Davey Smith G, Egger M. Meta-analysis. Unresolved issues and future developments. Bmj 1998; 316: 221-5.
15. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000; 56: 455-63.
16. Sterne JA, Egger M. High false positive rate for trim and fill method. BMJ 2001; 320: 1574.
17. Delgado-Rodriguez M, Sillero-Arenas M. [Biases in meta-analysis]. Med Clin (Barc) 1999; 112 Suppl 1: 43-50.
18. Gibson L. GlaxoSmithKline to publish clinical trials after US lawsuit. Bmj 2004; 328: 1513.
19. Richy F, Bruyere O, Ethgen O et al. Structural and symptomatic efficacy of glucosamine and chondroitin in knee osteoarthritis: a comprehensive meta-analysis. Arch Intern Med 2003; 163: 1514-22.
20. Richy F, Ethgen O, Bruyere O et al. Efficacy of alphacalcidol and calcitriol in primary and corticosteroid-induced osteoporosis: a meta-analysis of their effects on bone mineral density and fracture rate. Osteoporos Int 2004.
21. Pagliaro L, D'Amico G, Puleo A. Meta-analysis as a source of evidence in gastroenterology: a critical approach. Ital J Gastroenterol Hepatol 1999; 31: 723-42.
22. Biljana M, Jelena M, Branislav J et al. Bias in meta-analysis and funnel plot asymmetry. Stud Health Technol Inform 1999; 68: 323-8.
23. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. Bmj 1997; 315: 1533-7.
24. Egger M, Smith GD. Bias in location and selection of studies. Bmj 1998; 316: 61-6.
25. Egger M, Smith GD. Meta-Analysis. Potentials and promise. Bmj 1997; 315: 1371-4.
26. Egger M, Smith GD. Misleading meta-analysis. Bmj 1995; 311: 753-4.
27. Steinbrook R. Public registration of clinical trials. N Engl J Med 2004; 351: 315-7.
28. Reynolds T. Researchers push for publication, registration of all clinical trials. J Natl Cancer Inst 2003; 95: 772-4.
29. Julian D. Meta-analysis and the meta-epidemiology of clinical research. Registration of trials should be required by editors and registering agencies. Bmj 1998; 316: 311.
30. Jay P, Wallace M. Compulsory registration of clinical trials: under-reporting is not an option. Bmj 2004; 329: 1044.
31. Petticrew M, Gilbody S, Sheldon TA. Relation between hostility and coronary heart disease. Evidence does not

support link. Bmj 1999; 319: 917-8.

Author Information

Florent Richy, MPH-PhD
Epidemiology and Health Economics - WHO Collaborating Center for the Public Health Aspects of Osteoarticular disorders, Department of Public Health, Unit of Public Health, University of Liège

Jean-Yves Reginster, MD-PhD
Epidemiology and Health Economics - WHO Collaborating Center for the Public Health Aspects of Osteoarticular disorders, Department of Public Health, Unit of Public Health, University of Liège