

Correspondence

Psychological Medicine, 38 (2008).
doi:10.1017/S003329170800353X

Research Letter

Bias in genetic association studies: effects of research location and resources

There is increasing concern that the genetic literature may be distorted by various biases, such as publication bias, which may lead to a misleading impression of the strength of evidence for a putative gene–disease association. Meta-analysis is one means by which a more accurate estimate of the strength of evidence for such association may be obtained, as well as offering a means by which potential biases may be identified (Munafo & Flint, 2004). Here we present evidence that the location where a study is conducted is associated with the degree to which it represents an over-estimate of the true effect size, as subsequently estimated using meta-analytical techniques.

A number of factors are likely to introduce bias into the literature, and contribute to the risk of false-positive results. These include publication bias (Ioannidis, 2006), longer time to publish for results which do not achieve statistical significance (Ioannidis, 1998), the trend for effect sizes to decrease with year of publication (Trikalinos *et al.* 2004), the poor predictive value of initial reports of genetic association (Ioannidis *et al.* 2001), the *post-hoc* study of further sub-groups defined by sex or environmental factors (Patsopoulos *et al.* 2007), the excess of results that fall just below the 0.05 α -level (Ioannidis & Trikalinos, 2007), and so on.

We used data from three meta-analytical reviews, relating to the *DRD2 TaqIA* polymorphism and alcoholism (Munafo *et al.* 2007), the *DRD2 TaqIA* polymorphism and cigarette smoking (Munafo *et al.* in press), and the *COMT Val¹⁵⁸/108Met* polymorphism and schizophrenia (Munafo *et al.* 2005), resulting in a total of $k=81$ studies. For these, additional data on the geographical location of the research group which conducted each study (defined as the postal address of the corresponding author, grouped as North America, Europe, Other) were then extracted independently by two authors (M.R.M. and A.S.A.). Russia was coded as Europe, and Turkey as Other. We divided the individual study odds ratio (OR) by the pooled OR, to arrive at an estimate of the degree to which each individual study over- or under-estimated the true effect size, as estimated in the corresponding meta-analysis.

Data were combined within a fixed-effects framework using inverse variance methods, as described in detail elsewhere (Munafo & Flint, 2004). For studies coded as Europe [OR 0.96, 95% confidence interval (CI) 0.88–1.04, $p=0.30$] and Other (OR 0.95, 95% CI 0.89–1.01, $p=0.10$) these pooled OR did not differ significantly from zero. In both cases there was modest between-study heterogeneity ($I^2=34.66$ and 34.60 respectively). For studies coded as North America, however, there was evidence of a significant over-estimation of the true effect size (OR 1.10, 95% CI 1.02–1.17, $p=0.009$), with evidence of substantial between-study heterogeneity [$\chi^2(33)=134.33$, $p<0.001$, $I^2=75.43$]. Meta-regression indicated a significant negative correlation between year of publication and ln OR for studies coded as North America (slope -0.04 , $p<0.001$), but not those coded as Europe or Other (slope $+0.00$ and $+0.01$, respectively, p values >0.55).

Our results indicate that studies published in North America may represent a relative over-estimate of the true effect size, compared to those published in Europe or elsewhere. Although this conclusion assumes that the pooled effect size arrived at using meta-analytical techniques represents the best-available estimate of the true effect size, this is exactly the rationale behind the use of such techniques (Munafo & Flint, 2004), which have become increasingly popular in recent years for confirming or refuting the evidence for specific gene–disease associations. It is not possible from these data to infer why this might be the case, although a number of possibilities may be considered. These include, for example, the relatively greater research funding available in North America, and the competition between research groups for these resources, both of which may encourage a focus on the apparently most interesting and timely results available to them (i.e. those which reach nominal statistical significance, at the expense of those which do not), in order to achieve a high rate of publication in high-impact journals.

In order to investigate this possibility, we examined the average bias index for individual countries and the ratio of government research and development funding in science and technology to the gross domestic product (GDP) of these countries, where these data were available. This indicated a strong positive correlation ($r_s=+0.65$, $p=0.032$). These data are presented graphically in Fig. 1.

Clearly these and other biases will exist elsewhere as well, given that our data reflect a *relative* bias based

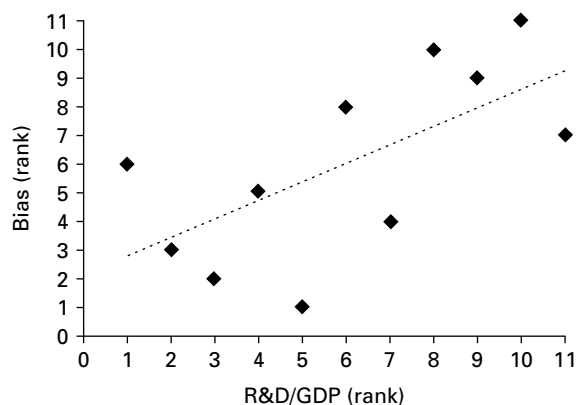


Fig. 1. Bias index and ratio of government research and development (R&D) funding to gross domestic product (GDP). Bias index, ranked in ascending order, is plotted against the ratio of government R&D funding in science and technology to GDP, also ranked in ascending order, for 11 individual countries. There is a strong, positive correlation ($r_s = +0.65$, $p = 0.032$), suggesting that greater research spending is associated with a greater degree of bias. (Source: OECD Main Science and Technology Indicators, 2006.)

on geographical location. One possible interpretation of these correlational data is that anything that distinguishes between North America and Europe will result in a similar correlation. However, by treating each country separately our correlational analysis is unlikely to be contaminated by any systematic bias that differentiates North America from Europe. Consistent with this interpretation, it should be noted that the USA and Canada are not placed at the top of the ratio of funding to GDP ranking of individual countries, although we cannot completely exclude the effect of a systematic bias. In conclusion, these findings should serve as a further caution that even our best available estimates of the strength of a genetic association will be distorted by publication and other biases.

Acknowledgements

Marcus Munafò is supported in part by a NARSAD Young Investigator Award. Jonathan Flint is supported by the Wellcome Trust.

Declaration of Interest

None.

References

- Ioannidis JP** (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *Journal of the American Medical Association* **279**, 281–286.
- Ioannidis JP** (2006). Journals should publish all ‘null’ results and should sparingly publish ‘positive’ results. *Cancer Epidemiology Biomarkers & Prevention* **15**, 186.
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG** (2001). Replication validity of genetic association studies. *Nature Genetics* **29**, 306–309.
- Ioannidis JP, Trikalinos TA** (2007). An exploratory test for an excess of significant findings. *Clinical Trials* **4**, 245–253.
- Munafò MR, Bowes L, Clark TG, Flint J** (2005). Lack of association of the COMT (Val158/108 Met) gene and schizophrenia: a meta-analysis of case-control studies. *Molecular Psychiatry* **10**, 765–770.
- Munafò MR, Flint J** (2004). Meta-analysis of genetic association studies. *Trends in Genetics* **20**, 439–444.
- Munafò MR, Matheson IJ, Flint J** (2007). Association of the DRD2 gene Taq1A polymorphism and alcoholism: a meta-analysis of case-control studies and evidence of publication bias. *Molecular Psychiatry* **12**, 454–461.
- Munafò MR, Timpson NJ, David SP, Ebrahim S, Lawlor DA** (in press). Association of the DRD2 Gene Taq1A polymorphism and smoking behaviour: a meta-analysis and new data. *Nicotine & Tobacco Research*.
- Patsopoulos NA, Tatsioni A, Ioannidis JP** (2007). Claims of sex differences: an empirical assessment in genetic associations. *Journal of the American Medical Association* **298**, 880–893.
- Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG, Ioannidis JP** (2004). Establishment of genetic associations for complex diseases is independent of early study findings. *European Journal of Human Genetics* **12**, 762–769.

MARCUS R. MUNAFÒ¹, ANGELA S. ATTWOOD¹ AND JONATHAN FLINT²

¹ Department of Experimental Psychology, University of Bristol, Bristol, UK; ² Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

Address correspondence to:
 Marcus R. Munafò, Ph.D.,
 Department of Experimental Psychology,
 University of Bristol,
 12a Priory Road, Bristol BS8 1TU, UK.
 (Email: marcus.munafò@bristol.ac.uk)