



UvA-DARE (Digital Academic Repository)

Prevention of false positive findings in observational studies: registration will not work but replication might

de Jonge, P.; Conradi, H.J.; Thombs, B.D.; Rosmalen, J.G.M.; Burger, H.; Ormel, J.

DOI

[10.1136/jech.2010.125252](https://doi.org/10.1136/jech.2010.125252)

Publication date

2011

Document Version

Final published version

Published in

Journal of Epidemiology and Community Health

[Link to publication](#)

Citation for published version (APA):

de Jonge, P., Conradi, H. J., Thombs, B. D., Rosmalen, J. G. M., Burger, H., & Ormel, J. (2011). Prevention of false positive findings in observational studies: registration will not work but replication might. *Journal of Epidemiology and Community Health*, *65*(2), 95-96. <https://doi.org/10.1136/jech.2010.125252>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Prevention of false positive findings in observational studies: registration will not work but replication might

P de Jonge,^{1,2} H J Conradi,^{1,3} B D Thombs,⁴
J G M Rosmalen,¹ H Burger,^{1,5} J Ormel¹

INTRODUCTION

Progress in science is built on a balance between curiosity and scepticism, and between creativity and rigour.¹ Although progress in science inevitably needs the generating of findings that may not be verified in subsequent studies, there is a considerable risk that exploratory studies introduce bias in the body of scientific knowledge. This risk may be substantial in observational studies, including cohort studies, case-control studies and cross-sectional studies.² Observational studies play an essential role in medical research as they are often conducted to evaluate research questions that cannot be addressed by clinical trials.³ As shown by Ioannidis, however, the likelihood that any given finding from a published observational study is true in reality is limited.² Aside from problems of uncontrolled confounding and other biases, observational studies are often generated from databases comprising many variables. Researchers often explore an unknown quantity of potential relationships capitalising on the chance of

obtaining positive findings. In this editorial, we will explore two potential solutions to reduce the risk of false positive findings from observational studies: a registration requirement; and a replication requirement.

REGISTRATION REQUIREMENT FOR OBSERVATIONAL RESEARCH

With a registration requirement, scientific journals would require researchers to register their observational studies in a manner similar to what has become policy for clinical trials.^{4,5} Study registration would include the recording of well grounded hypotheses that will be tested and data to be collected in a given study, recently suggested in the literature.⁶ Basic information with respect to the study would be documented, including design, variables used and, to the degree possible, a general theoretical framework. Existing registries, such as 'clinical trials.gov', could be used for this purpose, or perhaps new registers would need to be developed. Analyses not based on a priori hypotheses would not be precluded from publication, as exploratory research is essential for science to progress, but would need to be clearly labelled as exploratory.

There are some serious problems to overcome, however. Observational studies significantly outnumber clinical trials, and given that the largest clinical trials registry registers more than 200 trials weekly,⁵ coordination and maintenance would be substantial tasks. In addition, the number of analyses that need to be registered *per study* would be much larger than in clinical trials. Most importantly however, in observational research it is often impossible to have all specific analyses registered a priori. Hypotheses are often generated iteratively reflecting developments in the scientific literature or earlier findings within the cohort itself. Moreover, such

a procedure would still be vulnerable to manipulation once researchers have data in hand, because they could still engage in 'data-fishing'. In addition, an observational studies register might unfairly discriminate against legitimately generated hypotheses that are derived prior to inspecting the data but after data collection has begun. Finally, many researchers may not be willing to take the risk that their hypotheses might be 'stolen' by others when made public. Thus, a registration requirement of observational studies would not appear to be a simple solution, nor would it be fully effective in preventing the spread of false positive findings.

REPLICATION REQUIREMENT FOR OBSERVATIONAL RESEARCH

A different solution would emphasise the requirement that observational findings be replicated, a suggestion made in the literature as well.⁷ In this scenario, results from observational studies would be considered strictly exploratory (in the absence of a priori registration) *unless they are corroborated in an independent sample*. In addition, we propose that bootstrapping techniques become standard practice to correct the results for over-fitting of the models used. By doing so, the results are more likely to be reproduced in future populations. Bootstrapping techniques have shown to be superior to split-sample or cross-validation techniques in this respect.⁸

Similar to the criteria developed for replication studies in genotype-phenotype associations,⁹ replication studies using observational designs would need to (a) be of sufficient size, (b) use independent samples, (c) use a similar population, (d) demonstrate a similar magnitude and direction of effect, (e) include the same level of detail of the initial study, (f) use similar independent and dependent variables, and (g) include similar moderators and mediators as in the initial study. For confounders it seems more complicated. The overwhelming threat to the validity of observational studies, as contrasted with randomised studies, is the uncertain amount of uncontrolled confounding. As uncontrolled confounding can surely contribute to the risk of false-positive findings, a replication study should at least include the confounders adjusted for in the initial study, but preferably a larger set. By doing so, the test of replication may become more stringent than when only the confounders included in the initial study were controlled for.

¹ICPE (Interdisciplinary Center for Psychiatric Epidemiology), Department of Psychiatry and Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; ²CORPS (Centre of Research on Psychology and Somatic Disease), Department of Medical Psychology, Tilburg University, Tilburg, the Netherlands; ³Department of Clinical Psychology, University of Amsterdam, Amsterdam, the Netherlands; ⁴Departments of Psychiatry, Epidemiology, Biostatistics, and Occupational Health; and Medicine, McGill University, Montreal, Quebec, Canada; Department of Psychiatry, Center for Clinical Epidemiology and Community Studies, and Division of Rheumatology, Jewish General Hospital, Montreal, Quebec, Canada; ⁵Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

Correspondence to Prof dr Peter de Jonge, Interdisciplinary Center of Psychiatric Epidemiology, Department of Psychiatry, UMCG University of Groningen, CC72, PO Box 30 001, Groningen 9700 RB, the Netherlands; peter.de.jonge@med.umcg.nl

The replication requirement has its limitations too. First, it is plausible that a gentle variant of the 'proteus phenomenon'¹⁰ takes place, in which the magnitude of effect of the initial study depends on a true association along with a certain degree of chance. This will result in not entirely similar associations in the initial and replication study, with lower association estimates found in the latter. Yet, this problem of 'regression to the mean' may be overcome at least in part if the researchers of the initial study applied bootstrapping techniques to correct for over-fitting. Second, it is possible that no good replication sample is available. Third, a replication requirement would result in a more complex procedure for publishing scientific data. Fourth, mere replication might have the risk of replicating flaws in design such as incomplete correction for relevant confounders. In contrast to the previous solutions, however, we feel there are no *principal* problems with this approach to reducing the risk of false positive findings.

CONCLUSION

A replication requirement for observational research should by no means be regarded as an attempt to limit the creativity of researchers, but rather to attain some form of quality control in observational studies by clearly delineating the difference between hypothesis-driven and exploratory research. Similar statements have been made in the past,^{11 12} but

apparently never really followed by policy. The advantages of more quality control in observational studies are manifold. In addition to likely increasing the probability that findings from observational studies will be replicable, quality control would result in increased credibility for researchers adopting this guideline in the eyes of journals, their editors, reviewers, and eventually readers. Journals can play an important role in encouraging researchers and principal investigators to adopt quality control procedures for observational studies. Perhaps workgroups like the International Committee of Medical Journal Editors could formulate a position statement that would require researchers to follow this procedure by accepting papers as hypothesis-driven only when (a) a credible a priori statement of hypotheses and data to be analysed has been prepared (eg, from an observational studies registry or data application form), or (b) replication data is presented. Studies that do not meet either of these criteria could still be published, but only as clearly labelled exploratory analyses with full attention to inherent limitations and likelihood of replication of results.

Competing interests None declared.

Provenance and peer review Not commissioned; not externally peer reviewed.

Published Online First 26 November 2010

J Epidemiol Community Health 2011;**65**:95–96.
doi:10.1136/jech.2010.125252

REFERENCES

1. **Babak MA.** What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med* 2004;**66**:411–21.
2. **Ioannidis JPA.** Why most published research findings are false. *PLoS Med* 2005;**2**:e124.
3. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;**370**:1453–7.
4. Clinical Trial Registration: a statement from the International Committee of Medical Journal Editors. http://www.icmje.org/clin_trial.pdf.
5. **Laine C,** Horton R, DeAngelis CD, *et al.* Clinical trial registration: looking back and moving ahead. *JAMA* 2007;**298**:93–4.
6. **Hemingway H,** Riley R, Altman DG. Ten steps towards improving prognosis research. *BMJ* 2009;**339**:b4184.
7. **Ioannidis JPA,** Tzoulaki I. What makes a good predictor?: The evidence applied to coronary artery calcium score. *JAMA* 2010;**303**:1646–7.
8. **Harrell FE Jr,** Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361–87.
9. **Chanock SJ,** Manolio T, Boehnke M, *et al.* NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. *Nature* 2007;**447**:655–60.
10. **Ioannidis JP,** Trikalinos TA. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol* 2005;**58**:543–9.
11. **De Groot AD.** *Methodology. Foundations Of Inference And Research In The Behavioral Sciences.* The Hague-Paris: Mouton & Co, 1969.
12. **De Groot AD.** The theory of the science forum: subject and purport. *Methodology & Science* 1984;**17**:230–59.