Concerning systematic reviews and meta-analyses, and intervention studies, in the past 2 years guidance for drafting research protocols have been published (PRISMA-P and SPIRIT checklist, respectively), along with dedicated websites (http://www.crd.york.ac.uk/PROSPERO/ and http://www.clinicaltrials.gov, respectively). Though the endorsement by scientists of such guidance checklists is impressive, no one guidance for observational epidemiologic study protocols has been developed so far.

Though I acknowledge all the cons that have been raised and partly discussed in the reply of Jan P Vandenbroucke, I strongly believe that in the long run the advantages for the scientific community of pre-registration of study protocols will outweigh any potential disadvantage. In the meanwhile, coordinated efforts towards a funding system systems that truly award scientific research based on grounded epidemiological methods, on transparency, and on shared materials, and education efforts focused on journal editors, would contribute in the direction of improving the quality of scientific research.

References

Stefania Boccia
Section of Hygiene, Institute of Public Health, Università Cattolica del Sacro Cuore

Correspondence: Section of Hygiene, Institute of Public Health, Università Cattolica del Sacro Cuore, L. F. Vito, 1. 00168 Rome, Italy, Tel: 39-06-30154396/35001527, Fax: 39-06-35001522, e-mail: sboccia@rm.unicatt.it
doi:10.1093/eurpub/ckv119
Advance Access published on 2 July 2015

Preregistration: when shall we start the real discussion?

The debate about preregistration of protocols of epidemiologic studies had become an intellectually divisive issue, with strong ‘pro’ and ‘contra’ positions. The pro side, of which the commentary by Stefania Boccia in this issue of the journal is an example, is unflinching: all observational research needs to be preregistered to make it credible and manageable, in particular to root out the possibility of selection bias. The great exemplar is the randomized controlled trial (RCT). This pro side has never paid attention to the reasons why the contra side thinks that the need for registration and the need for sticking to the preregistered protocol may depend on the issue that is researched. The contra side does not tire to point out that when one tries to explain how Nature works, one is chasing a problem; when chasing a problem one tries all kinds of approaches, and it does not matter what the original thoughts were and how and why ideas have evolved. The only thing that matters is the final proposition to the scientific community: an explanation together with the corresponding data. The scientific community will then decide whether the ideas that are presented are useful as a start or a continuation of further expansion of knowledge. That is entirely different from the reasons for which preregistration was necessary for RCTs: the well-being of millions of patients across the globe may hinge on one or two RCTs (funded by very interested parties) and one does not want the whims of an investigator to define some subgroup or some alternative analysis after seeing the data. In contrast, exactly that kind of flexibility is necessary when chasing the solution of a problem of disease causation.

The contra camp has proposed that there are exceptions: ‘Consider a heated controversy concerning a topic with large societal and economic consequences, in which conflicting results have been obtained, perhaps even by analyses from the same data. To make progress, this might be an instance in which stakeholders sit together beforehand to agree on a protocol that will convince everybody. The main purpose of such actions is not “prespecification”; however, it is to bind stakeholders (who may distrust the other’s analyses) to a procedure they all trust’. To date, we have not heard of a proposition for differential application of preregistration from the pro camp.

The fundamental deficit of the pro position is that it only reasons in terms of methods and procedures of numerical research. The idea that checking a list of methods and procedures will lift us out of a problem is tantamount to Baron von Munchhausen’s attempt to lift himself out of a swamp by pulling at his own hair. Much like rising from a swamp demands levers from the outside, the solution of a problem of disease causation demands the integration of diverse types of scientific knowledge: numerical, clinical, pathophysiological and basic science. For this integration, there are no rules—it is done on a case by case basis. Sometimes methodology of numerical research plays an important role, at other times not. That is again very different from the RCT situation where a final decision has to be made based on a couple of numerical studies.

Bound to failure—or lead to fraud?

Eventually, the idea of universal registration of observational research is bound to practical failure. In an RCT, the moment of randomization is a clear dividing line: the data that are obtained after randomization are not looked at, except for interim safety analyses, and only analysed at the very end. In observational research, this dividing line is mostly non-existent. Databases have already been looked at to see whether the research question can be solved by the data; often the data are very complex and one has to learn their strengths and weaknesses when doing the analysis- or the failure of some analysis brings about ideas about a better way of progressing. The same almost always happens in systematic reviews, also of RCTs: those who draft a protocol for a systematic review already know several important RCTs (which are the data of the systematic review), and they know the reasons why the results of these studies led them to do a systematic review. To maintain that...
systematic review protocols are preregistered in the same sense as an RCT is a scam.

It is impossible to demand that if a new idea about an analysis arises during the analysis (and/or after seeing the data), one should ‘preregister’ this idea; it is equally impossible to demand that new data should always be collected for a new idea arising during analysis, and finally, it is unimportant how many and which analyses have already been done about a particular set of data. So, either researchers doing observational studies will not comply (except in the circumstances outlined above) or they will simply ‘fake’ that they preregistered their thoughts. The latter spectre has already been raised by one of the strong proponents of preregistration: that it may simply force people into scientific dishonesty—which is quite likely, because sticking to a predefined protocol is an idea that is alien to any explanatory science and not just to epidemiology. This was exemplified in a paper which the US chemist Whitesides wrote to guide his PhD students through their research undertakings: 'If you start the research to test one hypothesis, and decide, when you see what you have, that the data really seem to test some other hypothesis better, don’t worry. Write them both down and pick the best combinations of hypotheses, objectives, and data. Often the objectives of a paper when it is finished are different from those used to justify starting the work. Much of good science is opportunistic and revisionist'.

This reality of explanatory research should be matched by reflections of the ‘pro’ camp. Recently, the same pro proponent who feared forcing people into scientific dishonesty, concluded: ‘I would loathe seeing a “perfect” scientific literature where everything is preregistered, all checklist items are checked and papers are written by robotic automata before the research is conducted, but no real progress is made. We need to find ways to improve science without destroying it'. As a contra person, I could not agree more. However, these words were not translated into any proposal. When shall we start the real discussion?

Conflicts of interest: None declared.

References
2 Pearce N. Registration of protocols for observational research is unnecessary and would do more harm than good. Occup Environ Med 2011;68:86-8.
3 Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1:43-6.

Jan P. Vandenbroucke
Department Clinical Epidemiology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

Correspondence: Jan P. Vandenbroucke, Department Clinical Epidemiology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands, Tel: +31-71-5264037, Fax: +31-71-5266994, e-mail: J.P.Vandenbroucke@lumc.nl
doi:10.1093/eurpub/ckv118
Advance Access published on 2 July 2015