4. Zambelli A, Prada GA, Fregoni V, Ponchio L, Sagrada P, Pavesi L. Erlotinib administration for advanced non-small cell lung cancer during the first 2 months of unrecognized pregnancy. Lung Cancer. 2008;60:455–7.
5. Rivas G, Llinas N, Bonilla C, Rubiano J, Cuello J, Arango N. Use of erlotinib throughout pregnancy: a case-report of a patient with metastatic lung adenocarcinoma. Lung Cancer. 2012;77:469–72.
6. Lee CH, Liam CK, Pang YK, Chua KT, Lim BK, Lai NL. Successful pregnancy with epidermal growth factor receptor tyrosine kinase inhibitor treatment of metastatic lung adenocarcinoma presenting with respiratory failure. Lung Cancer. 2011;74:349–51.
7. Eliesen GAM, van den Broek P, van den Heuvel JJ, Bilos A, Pertijs J, van Drongelen J, et al. Editor's highlight: placental disposition and effects of crizotinib: an ex vivo study in the isolated dual-side perfused human cotyledon. Toxicol Sci. 2017;157:500–9.

E. Padrão[a,*], C. Melo[b], G. Fernandes[a,c], A. Magalhães[a]

[a] *Pulmonology Department, Centro Hospitalar de São João, Porto, Portugal*
[b] *Department of Obstetrics and Gynecology, Centro Hospitalar de São João, Porto, Portugal*
[c] *Faculty of Medicine, University of Porto, Porto, Portugal*

* Corresponding author.
E-mail address: eva.padrao@gmail.com (E. Padrão).

https://doi.org/10.1016/j.pulmoe.2018.03.007
2531-0437/

p-Hacking – A call for ethics

When the publication of scientific studies is influenced by the use and misuse of *p*-value[1] statistics two types of bias may occur – publication bias and inflation bias, also known as *p-hacking*.[2] The publication bias consists in considering only studies that present statistically significant results (i.e. $p < .05$). This bias removes from the literature studies whose results are considered negative, including false negatives. On the other hand, *p-hacking* consists in the exhaustive exploitation of data through the use of different analytical models and/or the manipulation of application criteria of these models until statistically significant results are obtained. While publication bias removes from the literature true or false negatives the *p-hacking* brings to the literature true or false positives. Conditioned literature (i.e. the absence of false negatives and the presence of false positives) will bias the results of secondary studies aiming to synthesise scientific evidence, such as meta-analyses, that inform clinical guidelines and evidence-based decision making.[3]

The demand for the statistically significant output (viz. $p < .05$) encourages researchers to do almost everything to achieve this result. There are a number of approaches[4,5] (e.g. the exclusion of univariate and/or multivariate outliers, the selection of independent variables (IV) through stepwise hierarchical models, the strategic withdrawal of IV in multiple models, dichotomizing ordinal or continuous variables) and all are legitimate, from a strictly analytical point of view, to obtain results where *p*-value is <.05.[6] The validity of the reported conclusions drawn by these methods is what is questionable, from a scientific point of view, given that there is a strong possibility of these results representing false positives, in other words, they may be mere statistical artefacts.[3]

The *p-hacking* bias is difficult to detect and it cannot be easily eradicated.[3] Many researchers do not perceive it as a real problem, either because of lack of knowledge or because of the incentives and pressure to publish statistically significant results.

The magnitude of the bias for the use of *p-hacking* is not yet established, however, it is estimated to be quite high.[3] Seokyung Hahn analysed the consistency between the analyses reported in the research protocol and the analyses reported in the study publication after completion from a local research ethics committee and found that only 53% mentioned an analysis plan and of these 88% did not comply with the protocol and could be the result of *p-hacking* practices.[7]

The pre-specification of the statistical analyses to be performed is one way of minimising the problem. Several studies follow an exploratory analytical approach which makes this pre-specification impossible. In addition, registration of health research protocols is not yet mandatory for all methodological designs. However, the evaluation by a health ethics committee of research protocols is already a widespread and successful practice in Portugal and across European Countries.[8]

It would be appropriate for the research protocols submitted to health ethics committees to describe in detail their analytical plan. That is, not merely stating the data that will be analysed with any particular software but rather the identification of: the analytical statistic(s) to be applied; the independent, dependent and concomitant variable(s) to be tested; the outlier definition and criteria; the post hoc tests that will be considered in the statistical modelling.

This pre-specification would make it possible to link the research statistical outputs to previous planning and prevent the negative effects of *p-hacking*. Additionally, it would be possible to develop more similar and replicable studies and to better assess the impact that *p-hacking* has on research. In the case of exploratory studies such detail is neither possible nor coherent.

Therefore, a call for health ethics committees to assess the manifestation of researchers' analytical intent in research protocols (i.e. pre-specified or exploratory) is pertinent to help prevent and further study the *p-hacking* bias.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The author has no conflicts of interest to disclose.

References

1. Goodman S. A dirty dozen: twelve *p*-value misconceptions. Semin Hematol. 2008;45:135–40.
2. Raj A, Patil S, Sarode S, Salameh Z. p-Hacking: a wake-up call for the scientific community. Sci Eng Ethics. 2017, http://dx.doi.org/10.1007/s11948-017-9984-1.
3. Head M, Holman L, Lanfear R, Kahn A, Jennions M. The extent and consequences of p-hacking in science. PLOS Biol. 2015;13:e1002106.
4. Gadbury G, Allison D. Inappropriate fiddling with statistical analyses to obtain a desirable *p*-value: tests to detect its presence in published literature. PLoS ONE. 2012;7, e46363.
5. Wicherts J, Veldkamp C, Augusteijn H, Bakker M, van Aert R, van Assen M. Degrees of freedom in planning, running. Analyzing, and reporting psychological studies: a checklist to avoid p-hacking. Front Psychol. 2016:7.
6. Wigboldus D, Dotsch R. Encourage playing with data and discourage questionable reporting practices. Psychometrika. 2015;81:27–32.
7. Hahn S, Williamson P, Hutton J. Investigation of within-study selective reporting in clinical research: follow-up of applications submitted to a local research ethics committee. J Eval Clin Pract. 2002;8:353–9.
8. Hedgecoe A. Research ethics committees in Europe: implementing the directive, respecting diversity. J Med Ethics. 2006;32:483–6.

Pedro M. Teixeira[a,b]

[a] *ICVS/3B's – PT Government Associate Laboratory, Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal*
[b] *Member of the Ethics Committee of the Administração Reginal de Saúde do Norte, ARS Norte, Portugal*
E-mail address: teixeira.pms@gmail.com

https://doi.org/10.1016/j.pulmoe.2018.03.001
2531-0437/