p-Hacking – A call for ethics

When the publication of scientific studies is influenced by the use and misuse of p-value statistics, two types of bias may occur – publication bias and inflation bias, also known as p-hacking. 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, 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.

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 (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, dichotomising ordinal or continuous variables) and all are legitimate, from a strictly analytical point of view, to obtain results where p-value is < .05. The validity of the reported conclusions drawn by these methods 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.

The p-hacking bias is difficult to detect and it cannot be easily eradicated. 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. 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.

The pre-specification of the statistical analyses to be performed is one way of minimising the bias. 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 widespread and successful practice in Portugal and across European Countries.

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


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 Regional 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/
© 2018 Sociedade Portuguesa de Pneumologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).