Naunyn-Schmiedeberg's Arch Pharmacol (2014) 387:1015–1016
DOI 10.1007/s00210-014-1044-7

EDITORIAL

How significant are your data? The need for a culture shift

Martin C. Michel

Received: 15 August 2014 / Accepted: 21 August 2014 / Published online: 30 August 2014

There is growing concern about a lack of reproducibility of findings from experimental studies in the life sciences. For example, investigators at Bayer Healthcare found only a quarter of 67 seminal studies to be reproducible (Prinz et al. 2011). Similarly, investigators at Amgen attempted to replicate 53 studies in basic cancer biology but were able to reproduce only 6 despite often cooperating with the original investigators (Begley and Ellis 2012). Many academic researchers have made the same experience, and lack of reproducibility apparently does not depend on the origin of the lab reporting the finding, i.e., academia vs. pharmaceutical industry, or the reputation of the journal publishing them. Mathematical models propose explanations why a large fraction of published results are false (Ioannidis 2005).

The lack of reproducibility of a relevant percentage of seminal findings is worrisome as it will lead to a major waste in time and resources for those basing their own research on published findings which may turn out to be false. In academia, it may derail the early career of young scientists who waste time chasing irreproducible findings and end up empty-handed. For more senior scientists, it may cause problems in obtaining future grants which can be particularly painful in times where research funding becomes ever harder to obtain. Where academic research is funded by governments or charities, it also constitutes a waste of taxpayer money or that donated for charitable purposes and, in the long run, will undermine the willingness of governments and charities to support research.

In the pharmaceutical industry, lack of reproducibility also leads to a waste of resources as invalid targets may be pursued, which further drives up the already escalating costs of drug discovery and development (Prinz et al. 2011). For example, a number of major drug companies have stopped or significantly reduced their investment into central nervous system drug research; in many cases, the main argument was that preclinical data failed to translate into clinical efficacy. Among many factors potentially implicated in the failed translation (including poor or no target engagement), irreproducible data are a major contributor. For example, there were a number of compounds reported to be effective in a mouse model of the amyotrophic lateral sclerosis but, when these findings were rigorously re-tested under very well-defined conditions, none of the original compounds showed any benefit (Scott et al. 2008).

Lack of reproducibility also raises ethical problems as it may lead to unnecessary use of experimental animals. On an even broader scale, it undermines the public trust in science. At times of an ever increasing complexity of our world, regaining trust in science is important for a rational discussion of societal decisions, e.g., related to climate change or the use of genetically modified organisms in the food chain. The problem of lack of reproducibility has gained attention for instance of the US National Institutes of Health (Collins and Tabak 2014) which are looking into new policies to yield more reproducible data. Even outside the research community, lack of reproducibility of scientific findings is seen as a major problem as highlighted by a recent title story in the magazine The Economist (Anonymous 2013).

It is generally felt that outright fraud accounts for only a minor fraction of irreproducible data. In contrast, insufficient understanding of statistical concepts and inappropriate use of statistical tests are likely to account for a major fraction of the problem. To address this problem, several pharmacology journals have published review articles attempting to educate the community on proper study design, data analysis, and reporting. Examples include a 2011–2013 virtual theme issue published in the British Journal of Pharmacology (http://

M. C. Michel (✉)
Department of Pharmacology, Johannes Gutenberg University, Obere Zahlbacher Str. 67, 55101 Mainz, Germany
e-mail: marmiche@uni-mainz.de

onlinelibrary.wiley.com/journal/10.1111/(ISSN)1476-5381/homepage/statistical_reporting.htm) or the January 2014 issue of Biochemical Pharmacology (http://www.sciencedirect.com/science/journal/00062952/87/1). This issue of Naunyn-Schmiedeberg's Archives of Pharmacology contains a related article by Dr. Harvey J. Motulsky (Motulsky 2014). Dr. Motulsky is the author of a statistics textbook (Motulsky 2013), but may be even better known as the creator of Graphpad Prism, a program commonly used by pharmacologists to analyze and graph data. While the article represents the personal opinion of Dr. Motulsky, it is simultaneously published by four different pharmacology journals coming from three different publishers, indicating that the editors of all of these journals and their publishers consider it an important step in the right direction.

Reading the Motulsky article, which I strongly recommend, may be shocking. All of his recommendations are based on sound and convincing analysis and make perfect sense once you read them; they are shared by other professionals in this field (Drummond and Vowler 2011; Marino 2014). Nonetheless, you will realize that hardly any of the studies previously published in leading pharmacology journals have adhered to them, i.e., contain flawed study design, data analysis, and/or reporting. Looking back, I must plead guilty that even very recent studies from my own lab apparently have been at least partly flawed in this regard (Böhmer et al. 2014; Igawa et al. 2012; Michel 2014; Michel-Reher and Michel 2013; Sand and Michel 2014).

Realizing this problem and its importance, several major pharmacology journals in collaboration with the International Union of Pharmacology are currently working on recommendations for proper study design, data analysis, and reporting of experimental data in the life sciences. We expect that these recommendations will be implemented in many pharmacology journals as part of the instructions to authors in early 2015. While these instructions will not be a simple copy of the Motulsky recommendations (Motulsky 2014), most likely many of them make their way into the new guideline. For the time being, I can only recommend to carefully look into the issues raised by Motulsky as well as in the abovementioned theme issues of the British Journal of Pharmacology and Biochemical Pharmacology. Properly using and reporting statistical methods and results will be a painful shift of culture for many but will lead to more reproducible (and correct!) finding in the studies you report and in those you read.

References

Anonymous (2013) Trouble at the lab. De Economist 409:23–27

Begley CG, Ellis LM (2012) Raise standards for preclinical cancer research. Nature 483:533

Böhmer T, Manicam C, Steege A, Michel MC, Pfeiffer N, Gericke A (2014) The α_{1B}-adrenoceptor subtype mediates adrenergic vasoconstriction in mouse retinal arterioles with damaged endothelium. Br J Pharmacol 171:3858–3867

Collins FS, Tabak LA (2014) Policy: NIH plans to enhance reproducibility. Nature 505:612–613

Drummond GB, Vowler SL (2011) Show the data, don't conceal them. Br J Pharmacol 163:208–210

Igawa Y, Schneider T, Yamazaki Y, Tatemichi S, Homma Y, Nishizawa O, Michel MC (2012) Functional investigation of β-adrenoceptors in human isolated detrusor focusing on the novel selective β3-adrenoceptor agonist KUC-7322. Naunyn Schmiedeberg's Arch Pharmacol 385:759–767

Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2:e124

Marino MJ (2014) The use and misuse of statistical methods in pharmacology research. Biochem Pharmacol 87:78–92

Michel MC (2014) Do β-adrenoceptor agonists induce homologous or heterologous desensitization in rat urinary bladder? Naunyn Schmiedeberg's Arch Pharmacol 387:215–224

Michel-Reher MB, Michel MC (2013) Agonist-induced desensitization of human β_3-adrenoceptors expressed in human embryonic kidney cells. Naunyn Schmiedeberg's Arch Pharmacol 386:843–851

Motulsky HJ (2013) Intuitive biostatistics, 3rd edn. Oxford University Press, Oxford

Motulsky H (2014) Common misconceptions about data analysis and statistics. Naunyn-Schmiedeberg's Arch Pharmacol. doi:10.1007/s00210-014-1037-6

Prinz F, Schlange T, Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 10:712–713

Sand C, Michel MC (2014) Bradykinin contracts rat urinary bladder largely independent of phospholipase C. J Pharmacol Exp Ther 348:25–31

Scott S, Kranz JE, Cole J, Linecum JM, Thompson J, Al-Nakhala BM, Vieira FG et al (2008) Design, power, and interpretation of studies in the standard murine model of ALS. Amyotroph Lateral Scler 9:4–15