Homework_Reproducibility

Joel Laia

10/01/2021

A call for tansparent reporting to optimize the predictive value of preclinical research

Problems highlighted in the paper:

"poor reporting, often associated with poor experimental design, is a significant issue across the life sciences",

"a core set of research parameters exist that should be addressed when reporting the results of animal experiments",

" a concerted effort by all stakeholders, including funding agencies and journals, will be necessary to disseminate and implement best reporting practices throughout the research community",

"inadequate experimental reporting can result in such studies being un-interpretable and difficult to reproduce",

"incomplete or inaccurate description of experimental design, especially how randomization of animals to the various test groups, group formulation and delineation of animal attrition",

"deficiencies in reporting key methodological parameters that can introduce bias", " lacked descriptions of crucial methodological information","e reviewers of these studies could not adequately identify potential limitations in the experimental design and/or data analysis"," inadequate reporting correlates with overstated findings"

"introduction of an unintentional difference between comparison groups"

"Choices made by investigators during the design, conduct and interpretation of experiments can introduce bias, resulting in false-positive results",

"with smaller studies, the positive predictive value is lower, and false-positive results can ensue, leading to the needless use of animals in subsequent studies that build upon the incorrect results2",

"Sample-size estimation","with smaller studies, the positive predictive value is lower, and false-positive results can ensue",

"Interim data analysis", " If the results are statistically significant in favour of the working hypothesis, the study is terminated and a paper is written. If the results look 'promising' but are not statistically significant, additional data are collected. This has been referred to as 'sampling to a foregone conclusion' and can lead to a high rate of false-positive findings"

"Retrospective primary end-point selection","to include or exclude specific animals on the basis of outcomes (for example, state of health, dissimilarity to other data) have the potential to influence the study result",

"Ad hoc exclusion of data","it is not uncommon for investigators to select a primary end point only after data analyses",

"Pseudo replictes","". There is a clear, but often misunderstood or misrepresented, distinction between technical and biologic replicates",

"A statistically significant result does not provide information on the magnitude of the effect and thus does not necessarily mean that the effect is robust, which could account for the poor reproducibility of certain studies",

Solutions proposed by the authors:

"authors of grant applications and scientific publications should report on randomization, blinding, sample-size estimation and the handling of all data",

"It is important to report whether the allocation, treatment and handling of animals were the same across study groups. The selection and source of control animals needs to be reported as well, including whether they are true littermates of the test groups",

". Investigators should also report on whether the individuals caring for the animals and conducting the experiments were blinded to the allocation sequence, blinded to group allocation and, whenever possible, whether the persons assessing, measuring or quantifying the experimental outcomes were blinded to the intervention.", "Randomization and Bliding and report how they were made",

"it is crucial to report how many animals were used per group and what statistical methods were used to determine this number.",

"Sample-size estimation; report how many animals were used per group and what statistical methods were used to determine this number",

"sample size and rules for stopping data collection should be defined in advance and properly reported",

"It is also important to report whether all animals that were entered into the experiment actually completed it, or whether they were removed, and if so, for what reason.","", it is important to report whether any data were removed before analysis and the reasons for this data exclusion.",

"reporting whether results were substantiated by repetition","reporting how often the particular experimen", ", carefully designed and powered animal studies should be budgeted for in the grant applications and funding agencies should consider supporting repetition studies where appropriate."

"report whether all animals that were entered into the experiment actually completed it, or whether they were removed, and if so, for what reason"

"specify a primary end point before the study is undertaken, the time(s) at which the end point will be assessed, and the method(s) of analysis"

"analysis of samples from multiple litters is essential",

"Small size effects","reporting whether results were substantiated by repetition and how often the particular experiment was performed",

"analysis of samples from multiple litters is essential",

"Improving the transparency and quality of reporting","Calling upon investigators to provide key information about the design, execution and analysis of animal experiments","t authors report whether and how their studies were carried out blind and randomized, how sample size was determined, whether data are missing owing to attrition or exclusion, and supply information about other important experimental parameter"

Bias in genetic association studies

Problems highlighted in the paper:

"small sample sizes used and the correspondingly low statistical power"

Solutions proposed by the authors:

" We divided the individual study odds ratio (OR) by the pooled OR, to arrive at an estimate of the degree to which each individual study over- or underestimated the true effect size"

Best Practices for Scientific Computing

Problems highlighted in the paper:

" many are unaware of tools and practices that would allow them to write more reliable and maintainable code with less effort",

", computing errors can have disproportionate impacts on the scientific process",

"scientists often can't know what their programs should do next until the current version has produced some results",

"many scientists type the same commands, waisting time and making mistakes"

" keeping track of changes (and being able to revert them if things go wrong), and collaborating on a program or dataset",

"code mistakes".

Solutions proposed by the authors:

"software developers must therefore take several aspects of human cognition into account: in particular, that human working memory is limited", "a program should not require its readers to hold more than a handful of facts in memory at once", "make names consistent, distinctive, and meaningful", "make code style and formatting consistent",

"Scientists writing software need to write code that both executes correctly and can be easily read and understood by other programmers", make the computer repeat tasks","Save recent commands in a file for re-use","use a build tool to automate workflows"

"In order to maximize reproducibility, everything needed to recreate the output should be recorded automatically in a format that other programs can read",

"work in small steps with frequent feedback and course correction",

"use a version control system", ""don't repeat yourself"","modularizing code allows people to remember its functionality as a single mental chunk","add assertions to programs to check their operation",

"add assertions to programs to check their operations"

"use an off-the-shelf unit testing library", "turn bugs into test cases", use a symbolic debugger","use a profiler to identify bottlenecks","write code in the highest-level language possible","y use an issue tracking tool"

Opinion: Reproducible research can still be wrong: Adopting a prevention apporoach

Problems highlighted in the paper:

confidence crisis of researchers worried about the rate at which studies are either reproducible or replicable",",there has been a crisis of confidence among researchers worried about the rate at which studies are eitherreproducible or replicable.",

"studies with ommitted variables, poor study design and missing data",

"omitted variables, poor study design, missing data"

"available raw data and statistical code and correct data analysis performed"

Solutions proposed by the authors:

"the raw data from the experiment are available, (ii) the statistical code and documentation to reproduce the analysis are available, and (iii) a correct data analysis must be performed.",

"the scientific community must ensure reproducibility and replicability",

"(i) increase the number of trained data analysts in the scientific community and (ii) identify statistical software and tools that can be shown to improve reproducibility and replicability of studies.".

Power failure: why small sample sizse undermines the reliability of neuroscience

Problems highlighted in the paper:

"researchers have strong incentives to engage in research practices that make their findings publishable quickly, even if those practices reduce the likelihood that the findings reflect a true (that is, non-null) effect",

"low power, by definition, means that the chance of discovering effects that are genuinely true is low. That is, low-powered studies produce more false negatives than high-powered studies"," the lower the power of a study, the lower the probability that an observed effect that passes the required threshold of claiming its discovery",

"a wide range of estimates of the magnitude of an effect", "vibration effects"

Solutions proposed by the authors:

" use data from meta-analyses"," Meta-analysis provides the best estimate of the true effect size"

"Perform an a priori power calculation",

"Disclose methods and findings transparently",

"Pre-register your study protocol and analysis plan",

"Make study materials and data available",

"Work collaboratively to increase power and replicate findings"

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

"'