European Neuropsychopharmacology (2019) 29, 1312–1320

ELSEVIER

www.elsevier.com/locate/euroneuro

REVIEW

Be positive about negatives-recommendations for the publication of negative (or null) results

Anton Bespalov[a,*], **Thomas Steckler**[b], **Phil Skolnick**[c]

[a] *Partnership for Assessment and Accreditation of Scientific Practice, Heidelberg, Germany*
[b] *Janssen Research & Development, Beerse, Belgium*
[c] *Opiant Pharmaceuticals, Inc., Santa Monica, CA, USA*

Received 9 August 2019; received in revised form 18 October 2019; accepted 23 October 2019

KEYWORDS
Good research practice;
Negative results;
Reproducibility;
Publication bias

Abstract
Both positive and negative (null or neutral) results are essential for the progress of science and its self-correcting nature. However, there is general reluctance to publish negative results, and this may be due a range of factors (e.g., the widely held perception that negative results are more difficult to publish, the preference to publish positive findings that are more likely to generate citations and funding for additional research). It is particularly challenging to disclose negative results that are not consistent with previously published positive data, especially if the initial publication appeared in a high impact journal. Ideally, there should be both incentives and support to reduce the costs associated with investing efforts into preparing publications with negative results. We describe here a set of criteria that can help scientists, reviewers and editors to publish technically sound, scientifically high-impact negative (or null) results originating from rigorously designed and executed studies. Proposed criteria emphasize the importance of collaborative efforts and communication among scientists (also including the authors of original publications with positive results).

1. Introduction

Over the past decade, robustness, reproducibility and translatability of preclinical research have become topics of discussion within the biomedical research community. Timely disclosure of negative or null results (e.g. in a conventional

* Corresponding author at: PAASP GmbH, Am Aukopf 14/1, D-69118 Heidelberg, Germany.
E-mail address: anton.bespalov@paasp.net (A. Bespalov).

https://doi.org/10.1016/j.euroneuro.2019.10.007

peer-reviewed scientific publication) is essential for understanding the strength of a working hypothesis and may provide an avenue to develop correcting strategies, if appropriate.

Calls to publish study results which do not support the initial hypothesis (negative, neutral or null results) have not yet produced the desired outcomes. Efforts by publishers and journal editors to encourage submission of negative findings are laudable but represent only one aspect of a solution that the scientific community must develop to stimulate dissemination of studies describing negative results.

Here, we present a proposal on how: a) scientists may be encouraged to identify negative results worthy of sharing with peers; b) to prepare and present these data in a form for peer review and publication. This proposal was successfully tested during the ECNP Preclinical Network Data Prize project (Supplementary Box 1).

Attitudes towards sharing negative results are slowly changing, an evolutionary process that may require years to fully develop. Therefore, the primary objective of this paper is to catalyze this process by making it more acceptable to both authors and journals. The authors strongly believe that well designed and executed "failures" can and should be disseminated to the scientific community.

The following discussion draws heavily on preclinical research based on the authors' expertise and experience. However, the authors hope that the recommendations in this paper will benefit both preclinical and clinical investigators.

2. The nature of negative or null results

There are three principal reasons why results are negative: 1) because the original hypothesis was incorrect; 2) failures to confirm the findings of a published report; and 3) technical causes – for example, by faulty reagents, inappropriate study design, and insufficient statistical power. In general, negative results which can be attributed to technical issues do not meet the threshold for dissemination. Technical failures also cover apparently negative results obtained in studies that were prematurely terminated because the scientists decided the original hypothesis appeared not to be confirmed (e.g. out of the motivation to save resources).

Here, we focus primarily on negative studies that represent true failures to confirm previously published results.

3. How often do studies fail to confirm the initial hypothesis or findings?

A low number of manuscript submissions with negative or null results may be attributed to either a lack of such studies or because this work typically remains unpublished. Although the latter possibility was highlighted by two seminal publications from Bayer (Prinz et al., 2011) and Amgen (Begley and Ellis, 2012), such failures to reproduce published work can and are sometimes relegated to the level of "anectodal examples" (Alberts et al., 2015).

Other than the rare cases when investigators have access to both published and unpublished data (e.g. Bath et al., 2009), there are several *indirect* approaches that reveal the existing bias towards reporting positive results. For example, Tsilidis and colleagues conducted a meta-analysis of 4445 sets of results in the CAMARADES (Collaborative Approach to Meta-analysis and Review of Animal Data in Experimental Studies) database and revealed that the observed rate of nominally significant results (38.7%) was nearly two times higher than the rate that would be expected based on the effect size of the most precise studies (i.e. those with the smallest standard errors) taken to estimate plausible effects (Tsilidis et al., 2013). Systematic reviews of the CAMARADES database further suggest that there are too many studies with statistically significant results, with a strong bias towards publications with positive results. Consistent with studies suggesting a strong publication bias towards positive results, meta-analyses indicate that up to 30% of studies are not published (Sena et al., 2010). One of the assumed reasons why these studies are not published is because of the negative nature of the results (Sena et al., 2010). In the absence of actual reporting of negative study results, it is not possible to assess the impact of reporting bias; current estimates may underreport the problem.

4. Why disclose negative or null results?

The most significant impact of reporting negative studies is the potential to reduce waste, discouraging the use of additional resources in an attempt to confirm questionable findings.

Publication of negative results by a trusted colleague or group of colleagues can result in adjusting one's own research plans. This is also one of the principal drivers for the need for prompt, independent replication and communication of studies appearing in high-profile journals. Such publications are, by definition, highly visible, and typically belong to "hot" areas of science. Once published, these studies serve as triggers for many scientists, who build their hypotheses and research plans based on these reports. This is the essence of scientific process, and publishing both confirming and non-confirming (negative) results make this process robust and secure.

The conduct and timely publication of both confirming and non-confirming studies is critical to a decision-making process that affects resource allocation in both the private and public sectors. For example, the nuclear retinoid X receptor agonist bexarotene was reported to markedly reduce (50%) amyloid plaque burden, significantly reducing the levels of soluble and insoluble brain Aβ and restoring cognitive and memory functions in amyloid precursor protein-presenilin 1 transgenic mice (Cramer et al., 2012). Several laboratories commented on difficulties confirming some of these claims (e.g., no effects on amyloid plaques; Balducci et al., 2015; Fitz et al., 2013; O'Hare et al., 2016; Price et al., 2013; Tesseur et al., 2013; Veeraraghavalu et al., 2013) but by the time the first of these reports were published (May 2013), patients were already being enrolled in a clinical trial (NCT01782742). The results of this trial are now published, and the primary outcome measure was negative: bexarotene did not affect either the composite or regional amyloid burden measured using the amyloid plaque-labeling PET ligand, florbetapir (Cummings et al., 2016). Would this trial have been initi-

Table 1 Key steps in deciding to disclose negative (null) results.

Steps	Explanation
Step 1. Negative? Can the observed pattern of results be presented as "negative"?	• Adequate data analysis methods establish the observed results as statistically negative (e.g. based on confidence intervals; see text for details)
Step 2. Negative. There is a high degree of confidence in the observed pattern of results.	• Technical failure excluded • Authors of original paper consulted (if possible) • Properly validated methods • High research rigor standards • Multiple readouts (if possible) • Multi-laboratory collaborative effort (if possible)
Step 3. Negative! The observed results are of value and need to be communicated!	• Full access to methods and raw data can be provided • Results are likely to have a significant impact if published • It may be argued that it is unethical to leave the results unpublished

ated if it were known that important elements of the preclinical results described in the original report could not be replicated?

Self-correction is essential for scientific progress, and this need for self-correction is repeatedly emphasized in the context of discussions on reproducibility (e.g., Alberts et al., 2015). Being a critical element of the self-correction process, properly conducted studies with negative results can be highly informative for defining critical aspects of study design, conduct, analysis and reporting. Thus, negative studies are expected to improve research because they highlight potential sources of bias in positive studies and suggest simple measures that can reduce the impact of such bias (Macleod et al., 2009), including sample size calculations, randomization, allocation concealment, blinded outcome assessments and reporting of potential conflicts of interest.

5. Deciding to publish negative results of a study

The decision to publish is fundamentally different for studies with positive and negative results. It is a common practice to report positive data when statistically significant results (e.g. *p*-value "<0.05") are obtained. For negative results, more stringent criteria should be met prior to publication. Here, we suggest three key steps that should be considered when preparing a manuscript based on negative results for publication in a scientific journal (Table 1; Fig. 1).

5.1. Are results really negative?

As the very first step, it needs to be confirmed that the results are indeed negative. A *p*-value >0.05 does not guarantee "negative" results. Strictly speaking, no experiment can provide an absolute proof of the absence of an effect. Rather, experiments serve to estimate the size and variability of an effect: if the variability is large, results may be inconclusive. If the variability is low, results could indicate a meaningful effect or a biologically trivial effect. To escape the binary decision framework (effect vs no effect) that is commonly assumed by following the statistical null hypothesis testing framework, it is proposed to report the effects (such as the difference between means) with the confidence intervals. If the confidence intervals are broad and cover values that are potentially relevant, then the results would not be "negative" and rather inconclusive and suggesting a need for further experimentation (Fig. 2). If the confidence intervals are narrow and cover only those effect sizes that are too small to be considered as relevant, then the results can be treated and reported as negative.

Decision on what effect sizes are relevant is not always easy. In some cases, there is a significant body of evidence that supports 'biological relevance' (e.g. for well characterized parameters such as blood pressure). In other cases, scientists have to rely on their knowledge and professional experience with particular methods, readouts, and processes. Attempts to confirm published work often follows high-impact research where effect sizes are large or very large. Follow-up studies would have a high impact not only if presented as "negative", but also by indicating that the effect size is much less than the originally reported (Ioannidis, 2008).

5.2. Excluding technical failures

Challenging the results of previously published research requires a high degree of confidence and should also not appear accusatory. If there is a failure to observe effects reported in an earlier publication, scientists should be prepared and willing to approach the authors of the original publication and ask for help (see Table 1). This help could include providing more detailed study protocols than the methods described in the publication, indicating sources and catalog numbers of specific tools and reagents, and including training in specific methods, asking for a material or reagent sample, and joint research efforts. Such contacts are not only a clear sign of "fair play", but are also essential for gaining confidence in one's own results. Describing Amgen's experience with failures to confirm published work, Begley and Ellis (2012) write: "… when findings could not be reproduced, an attempt was made to contact the original authors, discuss the discrepant findings, exchange reagents and repeat experiments under the authors' direction, occasionally even in the laboratory of the original investigator. These investigators were all competent, well-meaning scientists who truly wanted to make advances in cancer research" (p. 532). This open and frank communication dispels the notion that a claim of "failure to confirm" previously published results is not

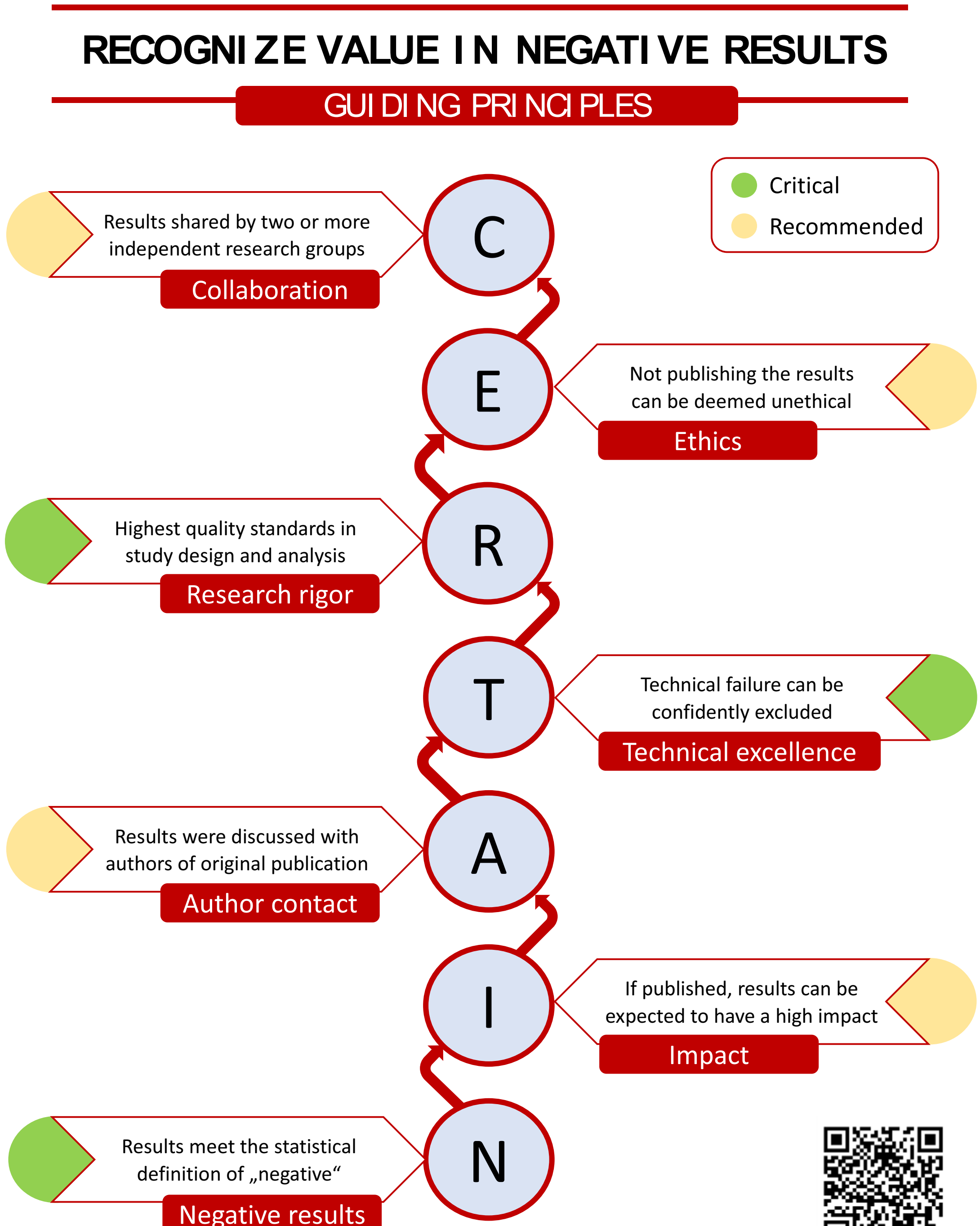

Fig. 1 Infographic introducing CERTAIN principles to support a decision on whether negative (or null) results are worth being published.

an *ad hominem* attack, and does not necessarily make the original positive results questionable or useless.

Communication with subject matter peers and experts is critical for excluding technical failure as an explanation for obtaining negative results. One of the most illustrative examples of the importance of such communication was provided by two independent laboratories, located in Boston and Berkeley, collaborating on establishing FACS profiles of primary mammary gland cells. Despite using identical methods and careful validation of all antibodies used, these labs obtained different CD44 expression profiles. Ultimately, the difference in outcomes stemmed from using different mechanical shaker speeds to disrupt tissue (Hines et al., 2014).

Not every technical detail can be reported in a paper, nor is it always possible to know which details are absolutely essential for comparing results from two different labs. It may happen that these efforts are successful, and a technical failure is excluded. However, certain methodological differences, unrelated to failure, may be identified as being responsible for discrepant results (e.g. (sub)strain of

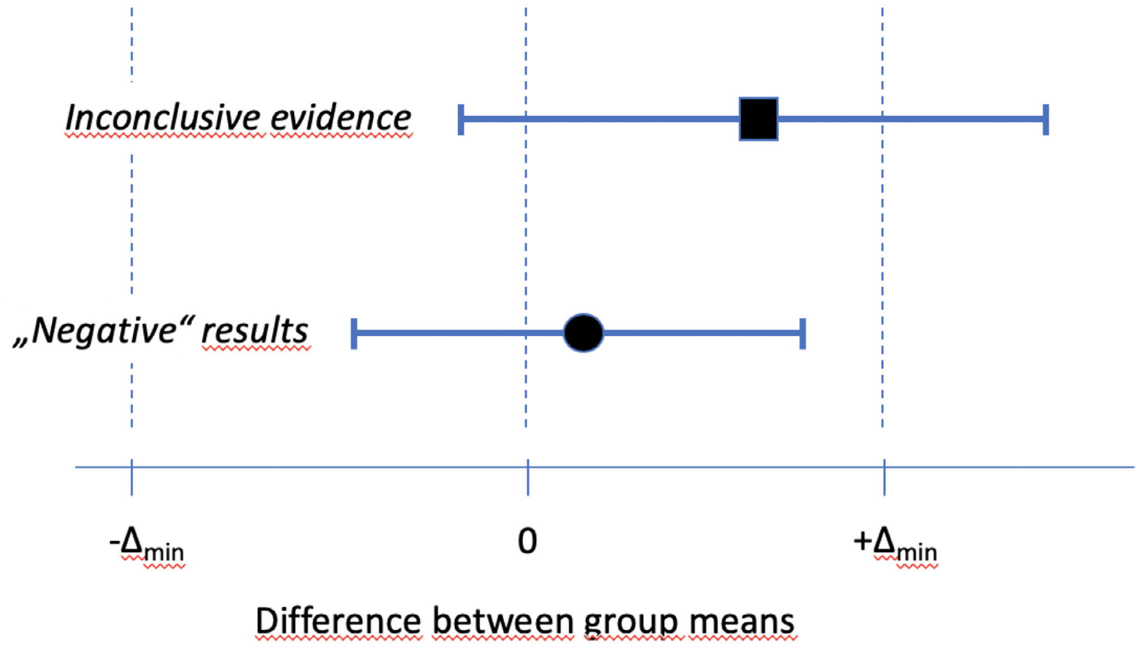

Fig. 2 Negative results vs inconclusive evidence. Mean differences between hypothetical Group A and Group B are presented with the hypothetical 95% confidence intervals. Δ_{min} is the minimum effect size (difference between group means) that is biologically meaningful or non-trivial. Inconclusive evidence - Groups A and B are not different and not equivalent. Negative results - Groups A and B are not different and equivalent.

animal used). In such cases, negative results may also merit publication.

An important aspect of a study's technical quality is validation of the experimental procedures and methods used. When negative results are obtained, scientists should be prepared to exclude a possibility that the methods were not sensitive to detect effects of experimental variables. This is commonly achieved by establishing calibration curves, dose-effect relationships with positive controls, etc. However, in some areas of science, validation of research methods can be problematic. Further, one may include positive results obtained using the same research methods (e.g. data generated with the positive controls) to reduce the likelihood that negative results are due to a technical failure or inadequate training.

5.3. Rigor in study design

In order to be credible, studies with negative results (and in particular follow-up, non-confirming studies) must be properly designed and powered. For publication, the follow-up study should, at a minimum, be designed with a high degree of rigor and statistical power and, in many cases, have a more rigorous design and more robust statistical power than the original report. It is often stated that the majority of research papers in the biomedical field are underpowered (e.g., Button et al., 2013) and, in most publications, no information is provided regarding sample size calculations (e.g., Macleod et al., 2015; Ramirez et al., 2017; Vogt et al., 2016). This is not surprising given that many published studies are exploratory, where no pre-existing data were available to support a proper power analysis. In this regard, follow-up replication efforts are advantaged because the study design should consider the effect size(s) reported in the original publication. To ensure this information is correctly interpreted and used, it is advisable to get support from biostatisticians.

If published, negative results may trigger a discussion as to whether the study is indeed sufficiently robust, the analysis applied correctly, etc. In order to address such concerns, the authors of a negative study should have and be ready to provide full access to the data and methods used. If the use of public repositories (http://www.nature.com/sdata/policies/repositories) is not possible, the authors may nevertheless explicitly state in the manuscripts their willingness to provide the data directly to potential reviewers or other peers interested in the data.

Generally speaking, considering the effort and cost of preparing and publishing a negative study, scientists should consider the above described criteria *before* conducting a study. If the follow-up study is run under enhanced research rigor standards aimed to eliminate bias (e.g. randomization, blinding, pre-specified hypothesis and data analysis, exclusion criteria, etc.; Dirnagl, 2016), a study protocol can be pre-registered using public online resources (e.g. www.preclinicaltrials.eu, https://cos.io/prereg/) or proprietary tools (e.g. electronic lab notebooks). This may increase confidence in the results and facilitate decision making by editors and reviewers.

5.4. Additional support for making a decision to publish negative results

The decision to publish negative results is often not an easy one for authors and editors. Adding to scientific and technical excellence, there are three further factors that may support a decision.

First, there is an ethical component and responsibility to be considered, especially when it comes to studies using animals. Scientists aiming to repeat an animal study should ensure that the data generated are impactful. Also, avoiding the unnecessary use of animals provides a compelling argument to publish negative results. Disclosure of negative results should be considered a very high priority if, due to the high scientific impact of the original study, multiple follow-up efforts are likely to be initiated. Scientists should not be discouraged by such "apparent" competition and the possibility that there could be disagreement among follow-up efforts (Supplementary Box 2).

Second, there must be no doubt about the potential value of these negative results and its impact on the research enterprise. For example, in a 2016 call for proposals to replicate cornerstone research (https://www.nwo.nl/en/research-and-results/programmes/replication+studies), the Netherlands Organisation for Scientific Research defined cornerstone research as: i) studies that are frequently cited with far-reaching consequences for subsequent research; ii) research that plays a major role in policy formation, or studies on the basis of which important policy decisions have been taken; iii) research that is part of the educational canon: research that is often cited in textbooks for students; iv) research that has received a lot of media attention and therefore has a considerable impact on the public debate; or v) studies with far-reaching consequences for legislation.

Third, collaborative efforts readily support modern multi-disciplinary research, are well recognized by funders and, according to recent analyses, result in higher impact publications (Wagner and Jonkers, 2017). Negative results that originate from collaborative efforts provide greater confidence and are also likely to be of higher value. But even when the results do not stem from a collaborative effort, it is worth discussing negative results with colleagues and peers, asking whether they had a similar experience. Such discussions can be especially fruitful in the case of negative results challenging a popular concept based on a previously reported set of positive results. If these earlier results were published in a high-profile journal, it is likely that the results are of sufficient interest for more than one lab to use these results to build on their own research (Cyranoski, 2017). Currently, there are no tools that enable these inquiries to go beyond an individual scientist's limited circle of colleagues. Information exchange with colleagues has two objectives. On the one hand, if these colleagues have unpublished *positive* (confirming the hypothesis) data, it makes sense to analyze various factors that may determine the study outcome (i.e. similar to the discussion with authors of original publications). On the other hand, if these colleagues also have unpublished *negative* (not confirming the hypothesis) results, they may be willing to contribute to a joint publication. Such shared publications will undoubtedly have a higher impact as the combined evidence is likely to be stronger. Combined evidence becomes stronger not only because of the mere amount of data but because of multiple independent attempts to investigate the potential cause(s) for the failure to confirm original findings (e.g. makes it more likely that there was no technical failure).

The decision to publish negative study results is ultimately based on the scientist's conclusion that, once published, these findings will be trusted and regarded as robust. Making this decision is not always easy, and support tools need to be developed and made accessible.

6. Where should negative results be published?

While there are many ways to disclose and disseminate negative results ranging from sharing them with colleagues, posting on the laboratory's website, to posting a comment on PubMed Commons, publication in a peer reviewed medium remains the gold standard. It is a common perception that journals are reluctant to publish negative studies. There is no empirical support for this perception and, therefore, one should consider submitting the negative study to a journal where the manuscript would be submitted if the results were positive (in the case of novel research), or if the study failed to replicate published results, to the journal where the original positive study was published. In our experience, journal editors can be easily approached and asked whether they would be willing to consider a negative study for publication.

In an ideal scenario, replication studies would be preregistered with the same journals publishing the original study, protocols reviewed by peers (including original authors) prior to conducting the study, and the resulting manuscript guaranteed publication provided the study was rigorous and technical failure is excluded. If such reports are published in an electronic *commentary* format, this would at least partially reduce the publishers and journals editors' fears of the potential decrease of the impact factor.

While publication of negative studies can be a frustrating process, we strongly advise scientists to avoid so called open-access "predatory" journals that charge a fee, essentially in exchange for expedited and often superficial review. A publication describing a negative study will be recognized and accepted by peers as an important piece of evidence only if it is published in a journal with standards that are consistent with journals publishing positive results. For avoidance of doubt, one may be advised to go for a journal that is *currently* listed on Web of Science and/or Scopus.

There are several additional factors that should be considered when selecting a journal for manuscript submission. For example, the original positive study may have been conducted with a single dose of a drug, while the follow-up experiments were conducted with several drug doses. Pharmacology journals traditionally favor proper analysis of dose-effect relationships and, therefore, may be a preferred venue to disclose results of the replication efforts in such cases.

7. Publishing negative studies: challenges and potential solutions

Scientists publish their results because it remains the most basic currency in academic science, integral for grant funding, professional advancement, and establishing collaborations. The publication of "failure to replicate" studies is not currently viewed as essential to achieving these goals. Many scientists view it as just the opposite – preparing a manuscript with negative results may be seen as waste of time. Indeed, negative studies are expected to be less cited (Fanelli, 2010; Misemer et al., 2016). Further, high-visibility papers, a frequent subject of replication efforts, are often complex, comprising a series of experiments using a variety of methods and tools. Replication efforts usually focus on a key experiment, chosen according to certain criteria (e.g. being central to the conclusions in the original publication, most straightforward to analyze and interpret, not requiring tools and reagents that are not readily available). Thus, negative replication results may challenge only one facet of the original publication's results. For example, papers citing the report describing the effects of bexarotene in Alzheimer's disease mouse models sometimes refer only to the effects of bexarotene on soluble Aß (Sweeney et al., 2017) and do not mention either its lack of effects on amyloid plaque load in the follow-up studies (Balducci et al., 2015; Fitz et al., 2013; O'Hare et al., 2016; Price et al., 2013; Tesseur et al., 2013; Veeraraghavalu et al., 2013) or the negative reports on its effects on soluble Aß (Price et al., 2013; Tesseur et al., 2013; see also Wang et al., 2016).

Visibility of the follow-up studies not confirming the hypothesis can be further enhanced by first preparing commentaries for either PubMed Commons or specialized discussion and information forums (e.g. Alzforum). Such commentaries could explicitly refer to the failed attempts to confirm the results of the original publication and, therefore, suggest the readers to look at both the original publication as well as follow-up studies. Reviewers, and ultimately both journal editors and publishers may also facilitate dissemination by requiring authors of new submissions to cite both positve and negative evidence. In case of a citation bias (one way or another), authors would then be required to publish an Erratum notice as when author names are mis-spelled or technical information is misindentified.

Besides low visibility, publication of negative results could lead to conflicts that have potential implications for a scientist's career. Thus, a negative study could be contradicting published results reported by the scientist's next grant reviewer, supervisor, or a key opinion leader(s) from an influential research institution. Perhaps the most expedient solution is to publish failure to replicate data as a collaboration, reducing the likelihood that this is a single laboratory's technical failure. This could be a single publication by authors from several institutions or papers published by multiple laboratories as a series, as in the bexarotene case.

Perhaps the greatest barrier to publishing negative studies is an inherent bias that views a negative study as a failure resulting from a lack of rigor and/or competence in designing and conducting experiments. This bias assumes that investigators reporting positive results are more technically skilled than those reporting negative results, although clearly this should be considered a misperception.

What can be done to help the reviewers and editors overcome this type of bias? There is a single answer to this question – negative studies should be technically rigorous, and the review process focused on study design and analysis as well as quality of tools and reagents. Scientists submitting manuscripts with negative results may facilitate the process by including special sections that explicitly rule out the probability of a technical failure, including positive and negative controls when feasible.

8. What needs to be done to facilitate disclosure of negative results?

The proposed criteria are aimed at facilitating the task to disclosing negative results. First, these criteria help scientists to identify results worth communicating. Second, the importance of collaboration is emphasized. Having coauthors may reduce the workload associated with preparing and publishing a manuscript, and also reduces the impact of potentially negative feedback triggered by the publication, especially in light of a possibly still existing negativity towards negative results (Matosin et al., 2014). Scientists need support and specific instruments that make it easier to decide to invest the time and effort in preparing and disclosing studies with negative results. We hope that the proposed criteria can provide a basis for authors, reviewers, and editors to decide which submissions describing negative results are worthy of publication.

Contributors

All authors contributed to and have approved the final manuscript.

Funding source

Apart from the ECNP Preclinical Network Data Prize sponsored by the Cohen Veterans Bioscience, there was no other funding or funding source contributing to preparation of this manuscript.

Declaration of Competing Interest

Anton Bespalov is a co-founder and/or employee of PAASP GmbH, EXCIVA GmbH and SYNVENTA LLC. Thomas Steckler is an employee of Janssen. Phil Skolnick is an employee of Opiant Pharmaceuticals.

Acknowledgments

This project is supported by funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement no 777364. This Joint Undertaking receives support from the European Union's Horizon2020 research and innovation programme and EFPIA.

The authors are very grateful to Cohen Veterans Bioscience for sponsoring the ECNP Preclinical Network Data Prize and all colleagues who provided support:

Advisory Board – Ijsbrand Jan Aalbersberg (The Netherlands), Celso Arango (Spain), Magali Haas (USA), C. Glenn Begley (Australia), Ulrich Dirnagl (Germany), Veronique Kiermer (USA), Malcolm Macleod (UK).

Research Rigor Experts – Emily Sena (UK), Kim Wever (The Netherlands)

Data Analysis Experts – David McArthur (USA), Martin Michel (Germany), Harvey Motulsky (USA)

Technical Quality and Scientific Excellence Reviewers – Daniela Alberati (Switzerland), William Banks (USA), Patrick Beardsley (USA), Yehezkel Ben-Ari (France), Pascal Bonaventure (Belgium), John Cryan (Ireland), Emmanuel Darcq (Canada), Taco de Vries (The Netherlands), Jan Deussing (Germany), Alexander Dityatev (Germany), Dario Doller (USA), Pim Drinkenburg (Belgium), Dagmar Ehrnhöfer (Germany), Christoph Emmerich (Germany), Nouchine Hadjikhani (USA), Guy Higgins (Canada), Matt Jones (UK), Mathias Jucker (Germany), Patricia Kabitzke (USA), Martien Kas (The Netherlands), Corinna Klein (Germany), Claire Mackie (Belgium), Malcolm Macleod (UK), Gavan McNally (Australia), Bettina Platt (UK), Mikhail Pletnikov (USA), Piotr Popik (Poland), Ana-Lucia Relo (Germany), Gernot Riedel (UK), Jan Rodriguez (Poland), Phil Skolnick (USA), Ioannis Sotiropoulos (Portugal), Michael Spedding (France), Thomas Steckler (Belgium), John Talpos (USA), Georg Terstappen (UK), Elizabeth van der Kam (Germany), Luc Ver Donck (Belgium), Karsten Wicke (Germany), Carsten Wotjak (Germany), Wolfgang Wurst (Germany)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.euroneuro.2019.10.007.

References

Alberts, B., Cicerone, R.J., Fienberg, S.E., Kamb, A., McNutt, M., Nerem, R.M., Schekman, R., Shiffrin, R., Stodden, V., Suresh, S., Zuber, M.T., Pope, B.K., Jamieson, K.H., 2015. Scientific integrity. Self-correction in science at work. Science 348, 1420–1422.

Balducci, C., Paladini, A., Micotti, E., Tolomeo, D., La Vitola, P., Grigoli, E., Richardson, J.C., Forloni, G., 2015. The continuing failure of bexarotene in Alzheimer's disease mice. J. Alzheimers Dis. 46, 471–482.

Bath, P.M.W., Gray, L.J., Bath, A.J.G., Buchan, A., Miyata, T., Green, A.Ron behalf of the NXY-059 Efficacy Meta-analysis in individual Animals with Stroke (NEMAS) investigators, 2009. Effects of NXY-059 in experimental stroke: an individual animal meta-analysis. Br. J. Pharmacol. 157, 1157–1171.

Begley, C.G., Ellis, L.M., 2012. Drug development: raise standards for preclinical cancer research. Nature 483, 531–533.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafo, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376.

Cramer, P.E., Cirrito, J.R., Wesson, D.W., Lee, C.Y., Karlo, J.C., Zinn, A.E., Casali, B.T., Restivo, J.L., Goebel, W.D., James, M.J., Brunden, K.R., Wilson, D.A., Landreth, G.E., 2012. ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models. Science 335, 1503–1506.

Cummings, J.L., Zhong, K., Kinney, J.W., Heaney, C., Moll-Tudla, J., Joshi, A., Pontecorvo, M., Devous, M., Tang, A., Bena, J., 2016. Double-blind, placebo-controlled, proof-of-concept trial of bexarotene Xin moderate Alzheimer's disease. Alzheimers Res. Ther. 8, 4.

Cyranoski, D., 2017. Authors retract controversial NgAgo gene-editing study. Nature doi:10.1038/nature.2017.22412.

Dirnagl, U., 2016. Thomas Willis lecture: is translational stroke research broken, and if so, how can we fix it? Stroke 47, 2148–2153.

Fanelli, D., 2010. Do pressures to publish increase scientists' bias? An empirical support from US states data. PLoS ONE 5, e10271.

Fitz, N.F., Cronican, A.A., Lefterov, I., Koldamova, R., 2013. Comment on "ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models''. Science 340 924-c.

Hines, W.C., Su, Y., Kuhn, I., Polyak, K., Bissell, M.J., 2014. Sorting out the FACS: a devil in the details. Cell Rep. 6, 779–781.

Ioannidis, J.P.A., 2008. Why most discovered true associations are inflated. Epidemiology 19, 640–648.

Macleod, M.R., Fisher, M., O'Collins, V., Sena, E.S., Dirnagl, U., Bath, P.M., Buchan, A., van der Worp, H.B., Traystman, R.J., Minematsu, K., Donnan, G.A., Howells, D.W., 2009. Preventing introduction of Bias at the Bench. J. Cereb. Blood Flow. Metab. 29, 221–223.

Macleod, M.R., Lawson McLean, A., Kyriakopoulou, A., Serghiou, S., de Wilde, A., Sherratt, N., et al., 2015. Risk of Bias in reports of in vivo research: a focus for improvement. PLoS Biol. 13 (10), e1002273.

Matosin, N., Frank, E., Engel, M., Lum, J.S., Newell, K.A., 2014. Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture. Dis. Model Mech. 7, 171–173.

Misemer, B.S., Platts-Mills, T.F., Jones, C.W., 2016. Citation bias favoring positive clinical trials of thrombolytics for acute ischemic stroke: a cross-sectional analysis. Trials 17, 473.

O'Hare, E., Jeggo, R., Kim, E.M., Barbour, B., Walczak, J.S., Palmer, P., Lyons, T., Page, D., Hanna, D., Meara, J.R., Spanswick, D., Guo, J.P., McGeer, E.G., McGeer, P.L., Hobson, P., 2016. Lack of support for bexarotene as a treatment for Alzheimer's disease. Neuropharmacology 100, 124–130.

Price, A.R., Xu, G., Siemienski, Z.B., Smithson, L.A., Borchelt, D.R., Golde, T.E., Felsenstein, K.M., 2013. Comment on "ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models''. Science 340 924-d.

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.

Ramirez, F.D., Motazedian, P., Jung, R.G., Di Santo, P., MacDonald, Z.D., Moreland, R., Simard, T., Clancy, A.A., Russo, J.J., Welch, V.A., Wells, G.A., Hibbert, B., 2017. Methodological rigor in preclinical cardiovascular studies: targets to enhance reproducibility and promote research translation. Circ. Res. 120, 1916–1926.

Sena, E.S., van der Worp, H.B., Bath, P.M., Howells, D.W., Macleod, M.R., 2010. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol. 8 (3), e1000344.

Sweeney, P., Park, H., Baumann, M., Dunlop, J., Frydman, J., Kopito, R., McCampbell, A., Leblanc, G., Venkateswaran, A., Nurmi, A., Hodgson, R., 2017. Protein misfolding in neurodegenerative diseases: implications and strategies. Transl. Neurodegener. 6, 6.

Tesseur, I., Lo, A.C., Roberfroid, A., Dietvorst, S., Van Broeck, B., Borgers, M., Gijsen, H., Moechars, D., Mercken, M., Kemp, J., D'Hooge, R., De Strooper, B., 2013. Comment on "ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models''. Science 340 924-e.

Tsilidis, K.K., Panagiotou, O.A., Sena, E.S., Aretouli, E., Evangelou, E., Howells, D.W., Al-Shahi Salman, R., Macleod, M.R., Ioannidis, J.P.A., 2013. Evaluation of excess significance bias in animal studies of neurological diseases. PLoS Biol. 11 (7), e1001609.

Veeraraghavalu, K., Zhang, C., Miller, S., Hefendehl, J.K., Rajapaksha, T.W., Ulrich, J., Jucker, M., Holtzman, D.M., Tanzi, R.E., Vassar, R., Sisodia, S.S., 2013. Comment on "ApoE-directed therapeutics rapidly clear β-amyloid and reverse deficits in AD mouse models''. Science 340 924-f.

Vogt, L., Reichlin, T.S., Nathues, C., Würbel, H., 2016. Authorization of animal experiments is based on confidence rather than evidence of scientific rigor. PLoS Biol 14 (12), e2000598.

Wagner, C.S., Jonkers, K., 2017. Open countries have strong science. Nature 550, 32–33.

Wang, S., Wen, P., Wood, S., 2016. Effect of LXR/RXR agonism on brain and CSF Aβ40 levels in rats. F1000 Research 5, 138.