Commentary

Biological Psychiatry: GOS

Emotional Impairments in Animal Models of Traumatic Neuropathic Pain: Where Do We Stand?

Filipa Pinto-Ribeiro

Preclinical studies have enjoyed a comfortable freedom in recent decades, with continuous technological advancements fostering the uncovering of mechanisms and intricate pathways that underlie the pathophysiology of many diseases. It is often claimed that they have the potential to translate findings effectively from the laboratory to real-world clinical applications. However, they are typically not subject to thorough preliminary review or assessment, and the rate of successful animal-to-human translation is low (1). Importantly, a growing number of publications have emphasized the urgent need to enhance the integrity of animal research in terms of its reliability, reproducibility, transparent reporting practices, data analysis, and the interpretation of results.

In this context, systematic reviews of preclinical studies serve as foundational tools for structuring preclinical research. By synthesizing and critically analyzing large amounts of experimental data (2), they counteract the lack of early scrutiny that might limit the applicability and relevance of its results. These reviews critically analyze the reproducibility and reliability of findings by highlighting methodological inconsistencies, sources of bias, and gaps in knowledge that potentially hinder effective translation to the clinics.

In pain research specifically, systematic reviews play an essential role by consolidating data on factors such as species, sex, age, and experimental paradigms (3). They guide researchers toward more standardized and robust methodologies that are not only internally rigorous but more translatable to human contexts. These reviews also contribute to refining animal models and thereby improving the overall quality and impact of preclinical pain research as a basis for the development of more effective therapies.

The recent systematic review by de la Rosa *et al.* (4) highlights the value of current rodent models of traumatic neuropathic pain for studying comorbid anxiety- and depressive-like behaviors. It shows that these animal models can effectively replicate, at least partially, the psychiatric components observed in human patients, although methodological limitations and inherent biological variability present considerable challenges.

The lack of uniform reporting standards referred to by the authors is a recurring complaint in systematic reviews, both clinical and preclinical. Despite the growing investment in developing reporting guidelines for each type of study, as is the case with the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines for in vivo animal studies (5), the problem persists. The blame for this does not belong solely to authors but also and especially to reviewers and to some degree editors and publishers.

The underreporting of crucial methodological variables, such as the total number of animals per group, habituation protocols, housing, and environmental conditions, may substantially affect behavioral outcomes. In parallel, the use of inappropriate controls, the misuse of statistical analyses, or the failure to comply with the correct reporting of statistical data can lead to misinterpretation of the significance of a finding. Scientific reporting will achieve higher standards only when the same level of attention is dedicated to methodology as is dedicated to the analysis of the results and when data generated by experiments are freely accessible (6).

As with most systematic reviews of animal studies, the work by de la Rosa *et al.* (4) highlights a sex bias, which is reflected in an overrepresentation of male animals, despite numerous articles having advocated for the inclusion of sex as a biological variable (7). Moreover, one of the more compelling findings is the difference in responses between male and female rodents, with neuropathically intact females exhibiting weaker anxiodepressive responses in tests considered to be the gold standard. This divergence raises questions about the validity of the behavioral paradigms used to evaluate anxiety- and depressive-like phenotypes in female animals. Importantly, it also questions the models' ability to recapitulate human sex differences in pain and associated emotional comorbidities because women experience a higher prevalence of both chronic pain and psychiatric disorders (8). Although the effects of ovariectomy were not addressed, these data should also be taken into account in experimental pain studies that simulate natural or surgical menopause.

Anxiety-like behavior evaluated using the elevated plus/zero maze, open field, and novelty suppressed feeding tests and depressive-like behavior through the forced swim, tail suspension, sucrose preference, and splash tests have consistently been observed in neuropathic animals. However, the variation in effect sizes between testing paradigms reinforces the importance of using multiple behavioral assessments to evaluate the different components of anxiodepressive-like behaviors. Together, these findings emphasize the need to refine and validate current protocols to ensure their effectiveness in detecting female-specific responses and to implement sex parity in future studies to improve the translational relevance of rodent models.

Among moderators, besides sex, species, strain, neuropathic model, and variations in model induction protocol all

SEE CORRESPONDING ARTICLE NO. 100388

https://doi.org/10.1016/j.bpsgos.2024.100424
ISSN: 2667-1743

influence outcomes differently depending on the behavioral paradigm being used. For example, anxiety-like behavior varies across rat and mouse strains when using the elevated plus/zero maze tests, and a similar effect in the open field and novelty suppressed feeding tests was related to the species being tested. Likewise, while strain differences in mice increased heterogeneity across all paradigms assessing the development of depressive-like behavior, the effects of rat strain were only significant when the sucrose preference test was used. Furthermore, while the experimental model of neuropathic pain had no effect on the evaluation of anxiety-like behaviors, it increased heterogeneity when evaluating depressive-like behavior using the tail suspension and sucrose preference tests. Data concerning the reliability of the forced swim test to assess the development of depressive-like behavior is particularly relevant given recent opposition to its use in preclinical research (9).

Interestingly, the width of the material used during model induction showed a significant moderator effect in the elevated plus/zero maze and the open field tests, but postinjury time did not. Age is a potential moderator that was not mentioned in the review; however, taking into account the low number of preclinical articles on pain and aging, it is probable that very few, if any, included old animals. These data emphasize the need for meticulous methodological planning and reporting and careful consideration of potential moderator effects because they impact the reliability and external validity of a study/model, as well as its translational relevance to human patients.

The authors also recommend refining the external validity of rodent models by incorporating/improving environmental enrichment, as well as social interaction and exposure to stress, which would reflect natural conditions and human clinical scenarios more accurately. Importantly, the authors found that the moderator being naïve for tests did not have an effect on the evaluation of anxiodepressive-like phenotypes. Nonetheless, without studying its association with rat and mice strain, the question remains whether repeated nociceptive stimulation over time is enough to decrease sensitivity thresholds in some animal models.

Finally, to enhance standardization and reduce bias, the authors suggest combining traditional evaluation methods with the adoption of automated tools such as the DeepLabCut and Bonsai for behavior assessments. These tools allow for unbiased, high-throughput data analysis, thus enhancing the reproducibility of behavioral data. However, they require the construction and optimization of frameworks and scripts tailored to each experimental paradigm and local setup. They also demand a certain degree of end-user coding and programming knowledge. Importantly, because they lack 3-dimensional analysis and precision as well as full integration of temporal/spatial information, their effectiveness in the detection of more complex behaviors is still limited.

In conclusion, the data of de la Rosa *et al.* (4) clearly demonstrate a high prevalence of anxiety-like and depressive-like behaviors in neuropathic animals. Differences related to species, strain, and sex call for a revaluation of existing experimental paradigms, particularly the inclusion of female groups. Revaluation would ensure better representation of the clinical population and of the validity of the tests used. The authors also emphasize the need for standardized methodologies and reporting practices, which will pave the way for the development of more effective therapeutic strategies by enhancing experimental rigor. They also stress the need for policies that promote comprehensive and inclusive experimental designs that accurately reflect human disorders.

Most importantly, this report serves as a model for researchers who are planning/conducting meta-analyses and systematic reviews of preclinical pain studies. It demonstrates how accounting for moderator variables can improve data interpretation and provide better recommendations for experimental designs in preclinical settings.

Acknowledgments and Disclosures

There was no direct financial support for this commentary.

In the past 3 years, FP-R has received funding from the Fundação Grünenthal.

Article Information

From the Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Campus of Gualtar, Braga, Portugal; and ICVS/3B's-PT Government Associate Laboratory, Braga/Guimarães, Portugal.

Address correspondence to Filipa Pinto-Ribeiro, Ph.D., at filiparibeiro@med.uminho.pt.

Received Nov 2, 2024; revised Nov 15, 2024; accepted Nov 16, 2024.

References

1. Ineichen BV, Furrer E, Grüninger SL, Zürrer WE, Macleod MR (2024): Analysis of animal-to-human translation shows that only 5% of animal-tested therapeutic interventions obtain regulatory approval for human applications. PLoS Biol 22:e3002667.
2. Ineichen BV, Held U, Salanti G, MacLeod MR, Wever KE (2024): Systematic review and meta-analysis of preclinical studies. Nat Rev Methods Primers 4:72.
3. Primo MJ, Fonseca-Rodrigues D, Almeida A, Teixeira PM, Pinto-Ribeiro F (2023): Sucrose preference test: A systematic review of protocols for the assessment of anhedonia in rodents. Eur Neuropsychopharmacol 77:80–92.
4. de la Rosa T, Llorca-Torralba M, Martínez-Cortés A, Romero-López-Alberca C, Berrocoso E (2024): A systematic review and meta-analysis of anxiety- and depressive-like behaviors in rodent models of neuropathic pain. Biol Psychiatry Glob Open Sci 4:100388.
5. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, *et al.* (2020): The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. J Cereb Blood Flow Metab 40:1769–1777.
6. Cunha-Oliveira T, Ioannidis JPA, Oliveira PJ (2024): Best practices for data management and sharing in experimental biomedical research. Physiol Rev 104:1387–1408.
7. Clayton JA, Collins FS (2014): Policy: NIH to balance sex in cell and animal studies. Nature 509:282–283.
8. Mogil JS, Parisien M, Esfahani SJ, Diatchenko L (2024): Sex differences in mechanisms of pain hypersensitivity. Neurosci Biobehav Rev 163:105749.
9. Molendijk ML, de Kloet ER (2022): Forced swim stressor: Trends in usage and mechanistic consideration. Eur J Neurosci 55:2813–2831.