Dear Editors,

Milesi et al. (2018, 2019) are to be applauded for their willingness to reanalyse the data used in their 2018 paper to take account of litter effects in their three-generation glyphosate study. As proposed by Plewis (2019), they now use appropriate statistical models (linear mixed models, often known as multilevel or hierarchical linear models) to take account of dependence between offspring from the same dam. This approach is more efficient than using litter as the unit of analysis as suggested by Paumgartten (2019) and others. Note that litter effects are just one way in which observational dependence is generated in rodent feeding experiments; cage effects and repeated measures are others. Unfortunately, the reanalyses of Milesi et al. (2019) do not go far enough in terms of allowing for the hierarchical structure of their data and assertions that their original findings hold up after the reanalyses are not borne out.

Consider first the model that lies behind Table 1 and using algebra rather than the confusing R syntax employed by the authors:

\[ y_{ij} = \beta_0 + \sum_{l=1}^{2} \beta_l t_{lj} + u_j + e_{ij}, \]

where \( y_{ij} \) is the pre-implantation loss % for each F1 rat; \( i = 1 \ldots n_j \) (max \( n_j = 4 \)) indexes the F1 rats within F0 dams \( j \); \( j = 1 \ldots 21 \); \( t_{lj} \) are dummy (0/1) fixed effects for the LD and HD treatments; \( u_j \sim N(0, \sigma_u^2) \) is the random effect for F0 with \( \sigma_u^2 \) being the variance between F0 dams and \( e_{ij} \sim N(0, \sigma_e^2) \), where \( \sigma_e^2 \) is the variance between F1 rats within the F0 dams.

The null hypothesis is that \( \beta_1 = \beta_2 = 0 \) and this should be tested using a single Wald test rather than using two separate tests as Milesi et al. (2019) do. Nevertheless, they do not find a statistically significant difference between the LD group and the controls and this is out of line with Fig. 3d (Milesi et al. 2018, p. 2635).

There are a number of points to note about this model and the data used by Milesi et al. (2019) to estimate its parameters. First is that the treatments vary only between F0 dams; all F1 offsprings are exposed to the same treatment. Second, the litter effect can be represented as the proportion of the total variance in the outcome explained by the F0 dams, i.e. \( \sigma_u^2/(\sigma_u^2 + \sigma_e^2) \), often known as the intra-cluster coefficient. Milesi et al. (2019) do not provide this estimate which is regrettable, because it would be a helpful indicator for the design of future studies even though a precise estimate of \( \sigma_u^2 \) might be difficult to obtain with such a small sample (21) of F0 dams. Third, there seems to be unexplained missing data as we would expect a F1 sample size of 28 for each of the three conditions, rather than 25, 20 and 20 as given and these numbers fall further at F2 to 20, 15 and 13. One implication of the way the F1 rats were apparently housed in cages (four per cage; Milesi et al. 2018, p. 2631) is that the variation in the outcome between F0 dams includes both a genetic effect and an environmental (or cage) effect arising...
from the co-housing. Ideally, these two effects should be separated by, for example, housing only two F1 rats together. Finally, the outcome variable, $y_{ij}$, is a percentage which might be very small for many F1 rats, suggesting that rather than assuming normality for $e_{ij}$, a binomial model might be more appropriate.

Turning to Table 2, we find that the authors do not fit the three-level models (F2 nested within F1 nested within F0) that are required, i.e.,

$$y_{ijk} = \beta_0 + \sum_{l=1}^{2} \beta_l t_{lk} + v_k + u_{jk} + e_{ijk},$$

where $y_{ijk}$ is now the outcome of interest (fetal weight etc.), for each F2 rat; $i = 1 \ldots n_{jk}$ indexes the F2 rats within F1 $j$, $j = 1 \ldots J_k$ and F0 $k$, $k = 1 \ldots K$; $t_{lk}$ are the treatment variables as before; $v_k \sim N(0, \sigma_v^2)$ is now the random effect for F0 dams; $u_{jk} \sim N(0, \sigma_u^2)$ is now the random effect for F1 (within F0) and $e_{ijk} \sim N(0, \sigma_e^2)$ with $\sigma_e^2$ the variance between F2 rats within both the F0 and F1 dams. Models such as these are easily estimated with the appropriate software and estimates of the variance components can inform future designs.

Even though Milesi et al. (2019) wrongly omit the F0 level from the model underpinning Table 2, their results are still different from those reported in the original paper: there is no longer any treatment effect for placental weight or for foetal length for the low dose. It is plausible to suppose that allowing for the full hierarchical structure would lead to a further reduction in the precision of the estimated treatment effects.

The authors are of course correct to state that no one experiment can be definitive. Arguably, more attention to experimental design is warranted to have appropriately powered multi-generational rodent experiments such as this one and others (e.g. Kubsad et al. 2019). But rather than offer their rats for further experiments, it would be much more helpful to researchers interested in estimating litter effects, if they made their data publicly available.

### Compliance with ethical standards

**Conflict of interest** The author declares that he has no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

### References

Kubsad D, Nilsson EE, King SE, Riggelman I, Beck D, Skinner MK (2019) Assessment of glyphosate induced epigenetic transgenerational inheritance of pathologies and sperm epimutations: generational toxicology. Sci Rep 9:6372. [https://doi.org/10.1038/s41598-019-42860-0](https://doi.org/10.1038/s41598-019-42860-0)

Milesi MM, Lorenz V, Pacini G, Repetti LD, Demonte LD, Varayoud J, Luque EH (2018) Perinatal exposure to a glyphosate-based herbicide impairs female reproductive outcomes and induces second-generation adverse effects in Wistar rats. Arch Toxicol 92:2629–2643. [https://doi.org/10.1007/s00204-018-2236-6](https://doi.org/10.1007/s00204-018-2236-6)

Milesi MM, Lorenz V, Beldomenico PM, Vaira S, Varayoud J, Luque EH (2019) Response to comments on: Perinatal exposure to a glyphosate-based herbicide impairs female reproductive outcomes and induces second-generation adverse effects in Wistar rats. Paumgartten FJ (2019) Comment on Perinatal exposure to a glyphosate-based herbicide impairs female reproductive outcomes and induces second-generation adverse effects in Wistar rats. Arch Toxicol 93:831–832

Plewis I (2019) Comment on perinatal exposure to a glyphosate-based herbicide impairs female reproductive outcomes and induces second-generation adverse effects in Wistar rats. Arch Toxicol 93:207. [https://doi.org/10.1007/s00204-018-2351-4](https://doi.org/10.1007/s00204-018-2351-4)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.