



Universiteit
Leiden
The Netherlands

Authors' reply: Who is afraid of non-normal data? Choosing between t-tests and non-parametric tests

Cessie, S. le; Goeman, J.J.; Dekkers, O.M.

Citation

Cessie, S. le, Goeman, J. J., & Dekkers, O. M. (2020). Authors' reply: Who is afraid of non-normal data? Choosing between t-tests and non-parametric tests, *183*(2), L5-L6.
doi:10.1530/EJE-20-0596

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3184713>

Note: To cite this publication please use the final published version (if applicable).

Authors' reply: Who is afraid of non-normal data? Choosing between *t*-tests and non-parametric tests

Saskia le Cessie^{1,2}, Jelle J Goeman² and Olaf M Dekkers¹

¹Department of Clinical Epidemiology and ²Department of Biomedical Data Sciences, Leiden University Medical Centre, Leiden, The Netherlands

Correspondence
should be addressed
to S le Cessie

Email
S.le_Cessie@lumc.nl

We would like to thank Dr Hoermann and colleagues for their comments (1) to our methodological editorial (2). We completely agree with them that one should always take the purpose of the data analysis into account and that variables which are skewedly distributed warrant special attention. However, we disagree with their recommendation that non-parametric tests should preferably be used for skewed variables.

To show our point, we consider the non-parametric testing in the example given by the authors. The Wilcoxon rank sum test yielded $W=30059$, $P=0.02$. How should a clinical reader interpret these results? The W value has no clinical interpretation and the P -value only tells that the difference is statistically significant. It does, however, not give any indication of the size or even the direction of the difference between the groups (3).

An important message of our editorial was that parametric methods yield valuable extra information besides a P -value. They provide an estimate of the mean difference (e.g. the difference in TSH between groups), with a confidence interval and the possibility to adjust for other variables using regression methods. We discussed in our editorial that, for very skewed data, parametric methods can still be used after performing a suitable transformation of the data. In many cases a logarithmic transformation will work well and will yield results which can still be easily interpreted by clinical readers. For example, a 1 unit larger mean of a 2-log transformed variable corresponds to a doubling of the median (geometric mean) on the original scale.

Finally, the authors mention that for some thyroid parameters, such as TSH, interest is not in average values but in the behavior of measurements in the tails. This is an argument to also provide raw data in a graphical form, which enables researchers to look at the tails of the actual distributions (e.g. to see whether some patients had very low or very high TSH values). And indeed, if the tails are of main interest, *t*-tests and (log)-linear regression may not be useful methods. However, the same holds for non-parametric testing methods, which are highly insensitive (by design) for detecting changes in behavior in the tails of distributions. Parametric methods can be tuned to be sensitive to changes in any parameter of interest, and if the tails are of interest, alternative parametric methods such as quantile regression may be preferred.

To conclude, we argue that parametric methods generally perform well and have the advantage to provide effect estimates which have an easier clinical interpretation than non-parametric methods. This, however, does not mean the researcher can be discharged to think carefully about which statistical method is most suited to answer the actual research question.

Declaration of interest

Olaf M Dekkers is Deputy Editor for *European Journal of Endocrinology*. He was not involved in the review or editorial process for this paper, on which he is listed as an author. The other authors have nothing to disclose.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

References

- 1 Hoermann R, Midgley JEM, Larisch R & Dietrich JW. Letter to the Editor Who is afraid of non-normal data? Choosing between parametric and non-parametric tests: a response. *European Journal of Endocrinology* **183** L1–L3. (<https://doi.org/10.1530/EJE-20-0134>)
- 2 Le Cessie S, Goeman JJ & Dekkers OM. Who is afraid of non-normal data? Choosing between parametric and non-parametric tests. *European Journal of Endocrinology* 2020 **182** E1–E3. (<https://doi.org/10.1530/EJE-19-0922>)
- 3 Dekkers OM. Why not to (over)emphasize statistical significance. *European Journal of Endocrinology* 2019 **181** E1–E2. (<https://doi.org/10.1530/EJE-19-0531>)

Received 3 June 2020

Accepted 8 June 2020