

Press release / 19 February 2018

EMBARGOED UNTIL THURSDAY 22 FEBRUARY 2018, 2 PM EASTERN TIME

Multi laboratory studies improve reproducibility of animal research

Pre-clinical animal research is typically based on single laboratory studies conducted under highly standardized conditions, a practice that is universally encouraged in animal science courses and textbooks. In a new study in PLOS Biology, researchers from the Universities of Bern and Edinburgh demonstrate that such insistence on uniformity risks producing results that are only valid under very specific conditions. In contrast, multi-laboratory studies that are based on diversity, substantially increased the reproducibility of animal experiments, which could help to further reduce the number of animals used for research.

The researchers used simulations based on 440 pre-clinical studies on 13 different experimental treatments in animal models of stroke, heart attack, and breast cancer, and could show that multi-laboratory studies can significantly improve the reproducibility of results. To simulate multi-laboratory studies, the researchers combined data from several individual studies, as if several laboratories had carried them out together. They found that the results of single-laboratory studies were very different, while multi-laboratory studies with two, three or four test laboratories, produced much more similar results, thereby increasing reproducibility without a need for larger sample sizes. "Our findings demonstrate, that standardisation in animal testing is an important cause of poor reproducibility, as it ignores biologically relevant variation," says lead author Prof. Hanno Würbel from the Animal Welfare Division at the University of Bern.

Single-laboratory studies produced no accurate estimate

In a first step, the researchers investigated 50 independent studies on the effect of therapeutic hypothermia (lowering of body temperature) on infarct volume, an indicator of stroke severity, in rodent models of stroke. A meta-analysis of these 50 studies showed that hypothermia reduced infarct volume by about 50% on average. The researchers took this effect as a yardstick to compare the accuracy and reproducibility of results from single-laboratory studies with those of multi-laboratory studies. More than half of the single-laboratory studies did not provide an accurate estimate of this effect. To simulate multi laboratory studies, two, three, or four studies were randomly selected from the pool of 50 studies, and from each of these studies proportionate numbers of data were collected, so that multi-laboratory and single-laboratory studies used exactly the same sample sizes. The percentage of studies that accurately estimated the expected effect of hypothermia increased from below 50% in single-laboratory studies to 73% in two-laboratory studies, to 83% in three-laboratory studies, and to even 87% in four-laboratory studies. "This

increase in the proportion of accurate study results with increasing numbers of laboratories reflects the improved reproducibility of results from multi laboratory studies," says co-author Dr. Bernhard Völkl, who carried out and analysed the simulations.

The accuracy of the results increased with the number of test laboratories involved

In a second step, the researchers repeated this analysis, using studies from twelve further experimental treatments in animal models of stroke, heart attack, and breast cancer, to determine whether their results can be generalised. In all cases, the accuracy and reproducibility of the results increased with an increasing number of participating laboratories. In addition, the researchers were able to show that an increase in sample size in single-laboratory studies cannot solve the problem; on the contrary: with larger sample sizes, the results become even less accurate.

These findings show that the standardisation of animal experiments is an important cause of poor reproducibility of results in preclinical animal research. Poor reproducibility questions the benefit of animal experiments and requires more replicate experiments – and therefore overall more animals – to answer a given research question conclusively. This contradicts the legal principle, according to which no unnecessary animal testing should be carried out, and the least possible number of animals should be used. "Our findings show that more representative test populations would improve the reproducibility of animal research," explains Hanno Würbel. And he adds: "Therefore, it would be possible to prevent using animals and other resources for inconclusive research." Accordingly, he recommends: "Multi-laboratory studies should replace standardised single-laboratory studies as the method of choice, at least in the later phases of pre-clinical trials." The fact that these improvements need neither large numbers of animals, nor lots of laboratories, could help with the implementation of these changes.

The study was funded by the European Research Council (ERC) and the Federal Food Safety and Veterinary Office (BLV – Bundesamt für Lebensmittelsicherheit und Veterinärwesen).

Publication:

Voelkl B, Vogt L, Sena ES, Würbel H (2018) Reproducibility of preclinical animal research improves with heterogeneity of study samples. PLoS Biol 16(2): e2003693.

<https://doi.org/10.1371/journal.pbio.2003693>

Contact:

Prof. Dr. Hanno Würbel

Head of the Animal Welfare Division, Veterinary Public Health Institute,

Vetsuisse Faculty, University of Bern

+41 31 631 25 30 / hanno.wuerbel@vetsuisse.unibe.ch