

TOOLBOX

New statistical method may spare lab animals

BY GRACE HUCKINS

18 MARCH 2021

Listen to this story:

A **new web-based statistical tool** could help researchers limit the number of control animals they use — by enabling them to repurpose data collected in other studies. The strategy could make it easier to detect traits that distinguish wildtype mice from those with mutations in autism-linked genes.

The approach leverages existing data from an experimenter's own lab or the broader literature to effectively increase the size of a study's control group. Though **not entirely new**, the method is available as an **online tool** for the first time. The team behind the tool described it and showed its effectiveness in *Nature Neuroscience* in February.

Researchers developed the online tool in response to an observation: Nearly every published mouse study is **underpowered**. Statistical power quantifies how likely an experiment is to uncover a particular effect — the impact of a gene mutation or a drug on social behavior, for example. Scientists often aim for a power of 80 percent — which means that they have an 80 percent chance of detecting an effect, if it exists — but few studies actually meet this bar.

Big effects can be detected with few animals, but smaller effects **demand larger numbers**. To make the problem explicit, the team reviewed animal studies on particular topics and found that the typical sample size — 10 per group — would achieve adequate power only for an unusually large effect. Less than 10 percent of the studies they reviewed discovered an effect of this magnitude.

More typical effect sizes require about 20 to 400 animals per group, the team found. And behavioral studies, in which effect sizes tend to be small, are at the upper end of this range. But

such large sample sizes create an ethical problem: Animals must be killed to complete many experiments, and even for those that survive, life in a laboratory is not ideal.

Downsizing:

To address these concerns, the team devised a method to reduce sample sizes while preserving adequate statistical power. They took advantage of Bayesian statistics, a framework that enables scientists to update their previous beliefs, or ‘priors,’ as new evidence becomes available. The tool enables scientists to build a prior from past studies and update it with their own data, giving them a final statistical distribution that describes the traits of control animals.

To demonstrate the benefits of their method, the team analyzed a dataset of nearly 150 mice that had experienced early-life adversity, and a similar number of controls. The two groups showed a statistically significant difference in their spatial learning abilities — which disappeared after the team removed two-thirds of the animals from the control group. When they applied their Bayesian method to bolster the control data, though, they rescued the result. In other words, their method made it possible to obtain the same significant result using about 80 fewer animals.

Despite its benefits, the method has some drawbacks: Historical control groups can differ from current ones in unexpected ways, particularly when the past data come from a different lab. To mitigate this problem, scientists using the tool can choose to weight individual data sources based on how similar they think those data are to their own control data. Experimenters should register those weights before they begin their study so that they cannot manipulate them later to get the results they want, the team says.

And as much as the approach is motivated by ethical concerns, its main goal is not to reduce sample sizes, say study investigator **Valeria Bonapersona**, a graduate student at Utrecht University in the Netherlands. Rather, she says, sample sizes must be increased to achieve adequate power — so she just wants to limit the magnitude of that increase as much as possible. “There’s basically no way around increasing the number of animals that we use,” she says.

Underpowered studies can miss real effects, Bonapersona says, so having larger sample sizes and historical control data could mean that fewer studies need to be conducted — and fewer animals harmed overall.