

# Animal instincts

By **Tim Fulmer**, Senior Writer

Experimental design and reporting requirements taken for granted in clinical trials are often lacking in preclinical research, making it difficult to predict the translational potential of an early stage finding. Now, a workshop convened by the NIH's **National Institute of Neurological Disorders and Stroke** has generated a set of standards for designing and reporting on animal studies.<sup>1</sup> The challenge will be to implement the standards throughout the research community and determine which preclinical studies should adhere to them.

In June, NINDS convened a two-day workshop entitled *Optimizing the Predictive Value of Preclinical Research*, which was attended by researchers, journal editors and grant reviewers with the express goal of developing preclinical reporting standards similar to those used in the clinical research community.

Participants included researchers from **Roche** and **Bayer AG**, grant reviewers from NINDS, editors from *The Journal of the American Medical Association*, *Nature*, *Nature Neuroscience*, *Science*, *Cell*, *Neuron* and *Neurology*, and researchers from a variety of universities, institutes and foundations, including the NIH's **National Center for Advancing Translational Sciences** (NCATS).

The resulting standards were published last month in *Nature* and addressed four areas the participants agreed were under-reported in grant applications and the peer-reviewed literature: randomization, blinding, sample-size estimation and data handling.

Many of the standards are obvious but surprisingly not part of common practice. For example, a review of 100 articles published in *Cancer Research* in 2010 found that only 28% of papers reported the use of randomization in animal studies, and just 2% of papers reported that the investigators were blinded during treatment.<sup>2</sup>

According to the new standards, researchers should randomly assign animals to experimental groups and should report the actual method of randomization. Moreover, researchers need to be blinded to what group a given animal is assigned, and the blinding should remain intact over the duration of the experiment.

Each animal study should have sample sizes that ensure sufficient statistical power to detect meaningful differences between groups, and the method of estimation should be reported.

Rules for stopping data collection, criteria for including and excluding data and all endpoints should be defined prospectively and reported.

Investigators also should report how often a particular experiment was performed and how well it repeated over a range of conditions.

Finally, participants discussed strategies for implementing the new standards.

The workshop said funding agencies and journals need to provide peer reviewers with a “minimum set of standards that should routinely be considered in evaluating the appropriateness of a study.”

Also, authors should be asked to provide information addressing the reporting guidelines on a standardized check-box form that accompanies manuscript submission. Such standardized forms are used by clinical research journals.

Other recommendations included encouraging investigators and journals to publish negative findings, creating a database for negative results and encouraging independent replication of studies.

## Standard application

The push to implement the new standards “should come from the key levers that control research behavior—publishers and funders,” said Elizabeth

Iorns, CEO of **Science Exchange**, a research service provider that links individual researchers with CROs. “A check-box solution at points of publication and grant submission would be the most obvious place to start.”

“Peer reviewers would have the primary role to ensure that methodological specifications are provided,” added Daniele Fanelli, a research fellow at **The University of Edinburgh** who has written about publication bias in the life and social sciences.<sup>3–5</sup> “Researchers would simply have to comply and, since adherence to the standards would become a mark of quality, most journals and institutions would eventually adopt the standards voluntarily.”

Earlier this year, Science Exchange and online publisher **PLOS** launched the Reproducibility

Initiative to help researchers carry out and publish the replication of preclinical translational experiments.<sup>6</sup>

Neither Iorns nor Fanelli participated in the NINDS workshop.

At first, journals should adopt the standards and post online guidance on how experiments should be done, said workshop participant Katrina Kelner, editor of *Science Translational Medicine*. Then, with increased awareness of the need for better reporting in the preclinical research community, it might be possible for journals to require that elements such as blinding and randomization be included in publications, she said.

## Not always applicable

Although introducing a set of experimental design and reporting standards to help better assess the translational value of animal studies would bring obvious benefits, careful thought will have to be given to what sort of preclinical experiments such standards should be applied to in the first place.

For example, very early stage observational experiments looking for

**“The key issue here is that translational research and basic biological research may make use of animal models in different ways. The experimental design required to show that a compound has a therapeutic effect in an animal model is likely different from an experimental design that uses an animal model to explore a basic biological process.”**

—**Kalyani Narasimhan**,  
Nature Neuroscience

any possible difference between groups of animals would be exempt from the standards. This hypothesis-generating work “is frequently conducted using a small sample size, does not have a primary outcome and is often unblinded” and is thus distinct from hypothesis-testing experiments, the authors wrote.

Even with the reporting standards agreed upon and in place, editors and reviewers will have to be cautious not to apply them indiscriminately to judge all preclinical research, said workshop participant Kalyani Narasimhan, chief editor of *Nature Neuroscience*.

“Many preclinical animal studies focus on basic biological phenomena and may not be designed to have their findings directly translated into drug discovery efforts,” said Narasimhan. “In some cases, for example, lack of blinding or randomization may not necessarily negate the findings and may not, by itself, be a reason for not publishing the paper.”

She added, “The key issue here is that translational research and basic biological research may make use of animal models in different ways. The experimental design required to show that a compound has a therapeutic effect in an animal model is likely different from an experimental design that uses an animal model to explore a basic biological process.”

Shai Silberberg acknowledged that “it is unrealistic to expect hypothesis-generating studies with no prespecified endpoints to meet all the proposed experimental standards, and certainly it is okay to publish those studies. Nonetheless, even in those cases, we expect the researchers to make clear to the reader that they used an exploratory experimental design, with perhaps a small number of animals and a lack of blinding and randomization.”

Silberberg is a program director at NINDS and was corresponding author on the *Nature* paper describing the workshop’s recommendations.

Purely practical matters may also make it difficult for standard academic labs to design experiments that meet the reporting standards, said Iorns. Many preclinical studies “are conducted by a single postdoc or grad student who designs and conducts the experiment and analyzes the data by themselves. They cannot necessarily be expected to do blinding, as it is only them conducting the research.”

The reporting standards and individual presentations from the June workshop are posted on the NINDS homepage.

### Looking beyond reporting

Poor reporting standards are only one part of the difficulties associated with translating published research.

“While promoting better reporting and better experimental design are obviously things we should strive to improve in scientific publications, those are not the only issues at the preclinical level responsible for poor translation into the clinic,” said Narasimhan. “Animal models of CNS conditions and other diseases are often inherently poor and often poorly predictive of human disease pathology.”

Indeed, two recent commentaries highlighted the myriad limitations of preclinical research programs.

In a commentary published this month in *Nature*, Jessica Bolker said the reliance of research biologists on a small handful of model organisms, such as the fly, mouse and worm, has significantly narrowed the types of hypotheses that can be accurately tested.<sup>7</sup>

If researchers use standard models that leave out “key causal elements such as environmental influences, we cannot hope to construct a complete picture of the mechanisms that underlie crucial variations, for example in development and disease,” wrote Bolker, who is associate professor of

zoology at the **University of New Hampshire**.

Thus, choosing a research model “should be more than a matter of convenience or convention,” wrote Bolker. “Scientists need to ask more questions—about the goals of a specific experiment, how suitable a given model is to reaching those goals and what environmental or other factors might be relevant to how well the model works.”

Bolker concluded her commentary by calling on NCATS to “support the development of new systems for investigating problems that are not tractable in currently favored models.”

In a commentary published in *Nature Reviews Drug Discovery*, three **AstraZeneca plc** researchers—Ian Peers, Peter Ceuppens and Chris Harbron—argued that “the systematic incorporation of expert statistical input into the design, analysis and interpretation of preclinical and translational research will help improve its quality, robustness and reproducibility.”<sup>8</sup>

Noting that a high level of statistical rigor is required in the clinical phases of drug development, the authors asked, “Why is it then considered appropriate to conduct preclinical research without insisting on the same level of statistical rigor and quality?”

Among the reasons for the lack of rigor, the authors cited limited regulatory oversight, the limited number of qualified preclinical statistical experts and a general lack of awareness among researchers of the value added by good statistical practice to preclinical work.

To help remedy the problem, the authors suggested the involvement of statisticians in preclinical research “be organized in a systematic way with clear roles and accountabilities, not on a ‘we’ll call you when needed’ basis, which is a common situation” in preclinical research.

Moreover, detailed statistical reviews should be incorporated in industrial governance processes and academic review processes “to set an expectation across the scientific community of the need to ensure that conclusions from data are justified,” according to the authors.

Fulmer, T. *SciBX* 5(44); doi:10.1038/scibx.2012.1152

Published online Nov. 8, 2012

### REFERENCES

- Landis, S.C. *et al. Nature*; published online Oct. 10, 2012; doi:10.1038/nature11556  
**Contact:** Shai D. Silberberg, National Institutes of Health, Bethesda, Md. e-mail: [silberbs@ninds.nih.gov](mailto:silberbs@ninds.nih.gov)
- Hess, K.R. *Cancer Res.* 71, 625 (2011)
- Fanelli, D. *Scientometrics* 90, 891–904 (2012)
- Fanelli, D. *PLoS ONE* 5, e10068; published online April 7, 2010; doi:10.1371/journal.pone.0010068
- Fanelli, D. *PLoS ONE* 5, e10271; published online April 21, 2010; doi:10.1371/journal.pone.0010271
- Fulmer, T. *SciBX* 5(34); doi:10.1038/scibx.2012.888
- Bolker, J. *Nature* 491, 31–33 (2012)
- Peers, I.S. *et al. Nat. Rev. Drug Discov.* 11, 733–734 (2012)

### COMPANIES AND INSTITUTIONS MENTIONED

**AstraZeneca plc** (LSE:AZN; NYSE:AZN), London, U.K.

**Bayer AG** (Xetra:BAYN), Leverkusen, Germany

**National Center for Advancing Translational Sciences**, Bethesda, Md.

**National Institute of Neurological Disorders and Stroke**, Bethesda, Md.

**National Institutes of Health**, Bethesda, Md.

**PLOS**, San Francisco, Calif.

**Roche** (SIX:ROG; OTCQX:RHHBY), Basel, Switzerland

**Science Exchange**, Palo Alto, Calif.

**The University of Edinburgh**, Edinburgh, U.K.

**University of New Hampshire**, Durham, N.H.