



Assessing the Reviewers of Animal Research

FOR THEIR ASSESSMENT OF THE RELIABILITY of Institutional Animal Care and Use Committee (IACUC) decisions, Scott Plous and Harold Herzog procured real animal use protocols and IACUC decisions, removed identifying information, and sent the protocols to other IACUCs to judge unofficially (Policy Forum, *Science's Compass*, "Reliability of protocol reviews for animal research," 27 Jul., p. 608).

On the basis of their data, the reliability among individual members within IACUCs appears low. However, U.S. Public Health Service policy requires that IACUC members have different roles (scientist, nonscientist, unaffiliated member, veterinarian). By design, members bring diverse views on a given protocol to the deliberations.

Agreement between the original and unofficial IACUCs also appears low. Of the 150 protocols, 141 (94%) were approved or approved pending modifications by the original IACUC, whereas the unofficial committee only rated 48% in these categories. This difference in judgments is the core of the authors' conclusion that IACUC decisions are unreliable.

That the masked protocols would be rated more negatively was predictable for the following reasons. First, IACUCs rely on

knowing the experience of the investigators and staff, information that was not included for the unofficial IACUCs. Not surprisingly, most of the negative shifts (84 of 118) were to categories calling for more information. Second, withholding approval had no practical consequence. Third, participants might have felt scrutinized by researchers with an "animal rights" agenda, and erred on the side of deferral or rejection. Fourth, navigating another institution's forms can be difficult. And fifth, IACUCs unfamiliar with

particular species or procedures are less likely to understand a protocol. These factors make it almost impossible to compare the actions of the original and unofficial IACUCs and thus call into question the major premises and conclusions of this study.

HARRY KLEMFUSS,^{1*} NANCY K. DESS,²
SUSAN E. BRANDON,³ HOWARD H. GARRISON,⁴
MARKY PITTS¹

¹Animal Subjects Program, University of California, San Diego, La Jolla, CA 92093-0071, USA.

²Department of Psychology, Occidental College, Los Angeles, CA 90041, USA. ³American Psychological Association, Washington, DC 20002-4242, USA. ⁴Federation of American Societies for Experimental Biology, Bethesda, MD 20814, USA

*To whom correspondence should be addressed.
E-mail: hklemfuss@ucsd.edu

IN THEIR STUDY OF DECISION-MAKING BY IACUCs, Plous and Herzog found great variation among and within different committees. In the authors' view, this variation casts doubt on the credibility of the IACUC review process and suggests a need for greater precision in evaluation criteria. We offer a different view of these results.

One of us (B.E.R.) was closely involved since 1976 in drafting what, in 1985, became the laws mandating IACUC ethical review. At that time, biomedical science—indeed, science in general—labored under a belief that scientific activity was value-free and ethics-free. The view that science does not make ethical judgments was so pervasive

that it essentially served as an ideological basis for scientific activity and, in particular, separated animal research from ethical deliberation and reflection. This, in turn, raised the question of how to ethically regulate animal research and analyze protocols from an ethical perspective.

One way would have been to create specific rules of the sort Plous and Herzog seem to favor. The fatal problem with this approach, we surmised at the time, was that such rules would most certainly have been

seen as arbitrary requirements imposed by regulators and would likely not have had any credibility or legitimacy. The alternative approach, which was adopted, was to provide such committees with only the broadest rules—for example, "control pain and distress"—to get them on the ethics playing field, and let each committee develop its own, more specific rules. In such a model, we hoped that local variations would eventually lead to consensus across the research community by virtue of shared discussions. And such consensus-building has indeed occurred regarding such issues as extending pain control to higher invertebrates or animals being used for ascites production, adjuvant use, and surgical research in an agricultural context.

Given that the IACUC system has only been in effect for 15 years and that society has only barely advanced beyond the anti-cruelty ethic for animals, IACUCs have done remarkably well dispelling ethical agnosticism among their members. There is every reason to believe, or at least hope, that, over time, ethical principles based on consensus will emerge and be embraced by the research community in general as the result of continued meaningful discussions.

BERNARD E. ROLLIN,¹ F. M. LOEW²

¹University Bioethicist for Colorado State University, Fort Collins, CO 80523, USA. ²President of Becker College, Worcester, MA 01609, USA

Response

IACUCs ARE REQUIRED BY LAW TO REJECT animal research protocols that do not comply with government regulations, just as Institutional Review Boards (IRBs) are required to reject human research protocols that violate government regulations. If a human research protocol were approved by one IRB only to be rejected by another (e.g., for lack of informed consent), most people would conclude that the regulations were not being enforced uniformly and that this unreliability should be addressed. In the case of animal research, however, Klemfuss and co-authors dismiss our findings as an inevitable result of IACUCs having a diverse membership, and Rollin and Loew dismiss the problem as a function of the IACUC system being only 15 years old.

IRBs also have diverse memberships and have not existed much longer than



Institutional Animal Care and Use Committees review animal research protocols to protect the animals from undue risk.

EXPLORE
ONLINE SHOPPING
LEARN
BUY
WWW.STEREOZOOM.COM
STEREOMICROSCOPY

- Explore • **New interactive website!**
- Learn • **New Stereo-microscopes**
- Buy • **Now online**

Leica

MICROSYSTEMS

IACUCs, yet they are not given an exemption from making reliable judgments. IRB lapses like those recently at Johns Hopkins University not only risk the lives of participants, they undermine public confidence in science.

Moreover, Klemfuss *et al.*'s methodological criticisms of our study are wrong on empirical and factual grounds. First, the reliability of IACUC protocol reviews does not increase significantly when the "diverse views" of veterinarians and unaffiliated members are eliminated from statistical analysis. Second, reliability does not vary by the species of animal used (half of the protocols involved rats, so it is unlikely that most IACUC members were unfamiliar with the species used). Third, Klemfuss *et al.* suggest that our study had an animal rights agenda, whereas, in fact, the study was endorsed by the Animal Behavior Society and was financially supported by two directorates of the National Science Foundation. Indeed, one of us (H.H.) is an animal researcher and has served on IACUCs for many years.

Klemfuss *et al.* point out that most of the negative shifts in opinion involved requests by the second committee for more information. What they do not mention is that 17 protocols were categorically disapproved (not simply "deferred") by the second committee, even though 16 of these protocols had been approved by the first committee. Equally striking, of the 72 protocols "approved as written" by the first committee, only 6 received that evaluation by the second committee. If, as Klemfuss *et al.* suggest, these shifts are attributable to factors such as the original committee's reliance on knowledge of the investigator rather than the written protocol—or to the inscrutable nature of particular protocol forms—these explanations provide all the more reason to reexamine the protocol review process.

Rollin and Loew object to "specific rules of the sort Plous and Herzog seem to favor," yet we did not propose any specific rule or animal care standard. What we advocate is not a proliferation of arbitrary regulations but the implementation of procedures to increase the reliability and validity of the review process, such as the development of explicit evaluative criteria, standardization and simplification of IACUC forms, and enhanced training of committee members.

Finally, it is worth noting that our results are not anomalous. They are consistent with previous research on unstructured peer review, including studies of IACUC and IRB decision-making, manuscript reviewing, and grant reviewing (1). Therefore, we ask the following question: At what point is the IACUC system

sufficiently well established, and the evidence of a problem sufficiently documented, for us to take action?

SCOTT PLOUS,^{1*} HAROLD HERZOG²

¹Department of Psychology, Wesleyan University, Middletown, CT 06549, USA. ²Department of Psychology, Western Carolina University, Cullowhee, NC 28723, USA

*To whom correspondence should be addressed.
E-mail: splous@wesleyan.edu

References and Notes

1. D. V. Cicchetti, *Behav. Brain Sci.* **14**, 119 (1991); R. Dresser, *J. Am. Vet. Med. Assoc.* **194**, 1184 (1989); J. Goldman, M. D. Katz, *J. Am. Med. Assoc.* **248**, 197 (1982).

WHO Ranking of Health Systems

BEFORE DEAN T. JAMISON AND MARTIN E. Sandbu's critique of the World Health Report 2000 (WHR2000) (Policy Forum, *Science's* Compass, "WHO ranking of health system performance," 31 Aug., p. 1595), there had been other criticisms of the report (1). Yet none of these commentaries discussed the fact that the World Health Organization (WHO) was recklessly inattentive to protection and prevention programs in their assessment of health system performance. In WHR2000 (2), WHO gives a comprehensive definition of a health system that includes "such traditional public health activities as health promotion and disease prevention, and other health-enhancing interventions like road and environmental safety improvement" (1, p. 5). However, WHO's actual assessment of health system performance pays scant attention to these activities that have historically contributed most to improving life and health. Instead, the report focuses almost exclusively on personal health care services—the equity in their distribution and the fairness in their financing. And, indeed, the recommendations in WHR2000 urge countries to improve service provision, resource generation, and health system financing for personal health care services.

As Jamison and Sandbu mention, the WHO rankings are intended to "lead to greater political accountability and to evidence-based health policies," but, that being the case, the focus of the report on personal health services poses a conundrum: If the WHO rankings of health system performance do not assess whether countries are taking advantage of a whole class of prevention activities—those that protect people from hazards in their living and working environments—then why should policy-makers and investors consider or adopt these strategies to improve the health of populations?

How can we correct the glaring omission of health-protecting activities from WHO's analysis and prescription? Jamison and Sandbu provide a hint: Concentrate on the method-