EDITORIAL

Getting the EPA back on track

Your information will be kept confidential, and the lessons learned from your participation will serve society—those are the promises made by researchers to participants in studies designed to inform environmental policies, from clean water and air to chemical exposure limits. The United States Environmental Protection Agency (EPA) may well break this fundamental pact next year, putting the agency at odds with its very mission "to protect human health and the environment." Hopefully, the EPA will realize that this would jeopardize regulations that keep the environment safe to live in, and correct course back to sound policy-making.

In January 2020, the EPA plans to issue a supplement to its 2018 proposed rule, Strengthening Transparency in Regulatory Science, which stated that in setting standards, the agency would only use research for which underlying raw data and models were made public. The rule could eliminate many public health studies from consideration. At a congressional hearing last month, the EPA claimed that the supplemental rule provides clarifications, but does it address major problems with the plan? Although the notion of depositing data and models from federally funded research into public databases is laudable, the rule as proposed poses substantial problems. This may account for why the majority of nearly 600,000 public responses to the 2018 proposed rule were critical.

"The EPA's proposed transparency rule... unquestionably excludes key science from policy-making."

In epidemiological and clinical studies, people provide information—their medical histories, behaviors, education, employment, and other personal details—under the condition that it will not be shared and their privacy will be protected. Anonymizing data is already difficult, if not impossible. With geographically referenced data, a capable programmer can leverage machine learning and brute computational strength to determine the location, and subsequently the identity, of a study participant. Similarly, facial recognition software has been applied to images reconstructed from cranial scans to identify study participants. Re-identification can jeopardize employment, insurance, or personal relationships for individuals, and scholarship, reputation, or funding for researchers. This will simply discourage people from participating in future health studies. Moreover, successfully recruiting and retaining participants depends on trusting relationships built on meaningful and sustained interaction between researchers and participants, especially with disadvantaged populations who are underrepresented in research. The EPA rule assumes that people will consent to their data residing in a repository where decisions about data use are made by persons unknown to them.

The proposed rule claims that additional analysis of raw data and models will improve science. Who will do this analysis? Most likely, vested interests will finance work slanted toward a particular outcome, rather than undertake scientific inquiry without an agenda. For example, lead paint industry defense attorneys have attributed children's neurological deficits to landlord neglect and parental failure. The rule also disregards the power of the "weight of the evidence." Imagine multiple studies done by different investigators on different populations using different techniques, yet reaching similar conclusions—that's a powerful result. Ignoring the weight of evidence derived from the totality of relevant science, regardless of data availability, contravenes the EPA's directive (stated in the Clean Air Act) to set standards "requisite to protect the public health" with "an adequate margin of safety."

Many researchers already deposit code and data into open repositories. The U.S. National Institutes of Health and other federal funding agencies require data-sharing plans to support independent reanalysis within the scientific community without compromising confidentiality. The peer review process provides an additional check on the credibility of research results. Work by the Health Effects Institute, in which an industry-government–funded partnership reanalyzed data from the Harvard Six Cities Study and the American Cancer Society Study on the link between particulate matter pollution and mortality, represents an excellent model for evaluating the validity of research pivotal to environmental health regulations without compromising confidentiality or excluding studies.

The EPA's proposed transparency rule does not ensure research rigor or improve transparency. It unquestionably excludes key science from policy-making. Once the supplemental rule is released in January 2020, there will be an open period for public comment—an opportunity for everyone to remind the EPA of its obligation to use the best science, as required in multiple environmental laws, to protect human health and the environment.

–Marie Lynn Miranda

PHOTO: THOMAS LAVERGNE

Marie Lynn Miranda is a professor in the Department of Statistics at Rice University, Houston, TX, USA, and is the founding director of the Children's Environmental Health Initiative at Rice University. mlm@rice.edu

10.1126/science.aba3769

Getting the EPA back on track

Marie Lynn Miranda

Science **366** (6470), 1173.
DOI: 10.1126/science.aba3769