



AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

732 North Washington Street, Alexandria, VA 22314-1943

(703) 684-1221 ■ www.amstat.org ■ asainfo@amstat.org

■ www.facebook.com/AmstatNews ■ www.twitter.com/AmstatNews

May 25, 2017

The Honorable Mike Rounds
Chair, Subcommittee on Superfund, Waste
Management, and Regulatory Oversight
Environment and Public Works Committee
U.S. Senate
Washington, DC 20510

The Honorable Kamala Harris
Ranking Member, Subcommittee on
Superfund, Waste Management, and
Regulatory Oversight
Environment and Public Works Committee
U.S. Senate
Washington, DC 20510

Dear Chairman Rounds and Ranking Member Harris,

As president of the American Statistical Association, with 19,000 members, I write regarding the House-passed H.R. 1430, The Honest and Open New EPA Science Treatment Act of 2017. We heartily support decisions being based on the best available science. We also applaud the idea that researchers and federal agencies strive to make data available to others—under strict pledges to maintain confidentiality of data provided by individuals and establishments where necessary—and to encourage reproducible research. Access to data and reproducibility of research are crucially important for science to advance.

While the bill's intent is to make data more widely available, we have several concerns and urge the bill be revised significantly in consultation of the scientific community before further consideration. Our concerns include those voiced by others in past Congresses (especially the American Association for the Advancement of Science) that the bill's statements do not account for the complexities common to the scientific process on research that involves biological materials or physical specimens not easily accessible, combinations of public and private data, longitudinal data collected over many years that are difficult to reproduce, and data from one-time events that cannot be replicated. The bill as written could have far-reaching consequences that would ultimately hamper or undermine the scientific process generally and EPA's work specifically.

Our nation should be striving for transparency in government and, as noted above, data accessibility, but these goals also must be balanced with the necessity to protect individuals' and businesses' privacy. The bill's language of "publicly available" except that "any personally identifiable information, trade secrets, or commercial or financial information obtained from a person and privileged or confidential, shall be redacted prior to public availability" acknowledges this balance; however, this language is insufficient to protect individuals and businesses as well-known cases of re-identification have made clear. Simple redaction of any

personally identifiable information further exemplifies our point that the bill fails to account for the complexities common to the scientific process.

Regarding the provision that “redacted information [...] shall be disclosed to a person only after such person signs a written confidentiality agreement with the Administrator, subject to guidance to be developed by the administrator,” we understand and appreciate that government policy to be based on analyses that can be replicated/evaluated by independent scientists. We cannot support the language of this bill for doing so. The government—including the EPA—routinely collects data such as private business information and private health information under strict pledges to protect confidentiality or on condition that it will not be shared. The pledges are enforced with strict penalties for violating them and it is essential to keep these promises to protect confidentiality. The bill lacks language on how confidentiality will be protected and under for what purposes the EPA should consider data sharing. These omissions and the related ambiguity could undermine EPA’s current and future data collection efforts. We urge careful consideration of the many conditions under which data are collected. We also urge close examination of the extensive protections in place for the sharing of data provided to federal statistical agencies which have decades of experience in protecting data confidentiality and whose data are collected under the 2002 Confidential Information Protection and Statistical Efficiency Act.

As an example of data collected for the EPA under confidentiality agreements and of the insights engagement with scientists could provide, consider an actual study relating to hydraulic fracturing and drinking water. Survey experts selected a random sample of wells and EPA collected information operating companies have on those wells. The companies provided the information on condition that it be considered confidential business information (CBI). While not the focus of the survey, statistical estimates from such data could lead to revised regulations but the individual well data were not to be made public. In fact it cannot be shared with EPA employees not working on the project and not specifically trained on the requirements of CBI under the Toxic Substances Control Act. Any legislation seeking to make such data available—under prescribed conditions to protect confidentiality to vetted scientists—should respect past data confidentiality protections while not undermining future data collections of this or a related nature.

Let me also point out the current EPA procedures include checks and balances designed to assure integrity and objectivity. All major rules have Notices of Proposed Rule Making and many also have Advanced Notices of Proposed Rule Making. These notices call for public comment, and the resultant comments must be substantively addressed. Further, rules of a scientific nature undergo review by the Science Advisory Board. Board members are area experts and are not EPA employees. Comments made by the Board must also be substantively addressed prior to a rule's promulgation. Dealing more directly with the data underlying EPA rules and procedures, EPA abides by the Information Quality Guidelines which prescribe comprehensive measures to assure quality, objectivity, utility, and integration of the data used in the rule.

Finally, let me also express our firm belief that science legislation should be bipartisan and formulated in consultation with the scientific community. By its very nature, science should transcend partisan lines. For science legislation, the engagement of the scientific community is imperative because of the complicated processes and considerations of science. Our statement of

the importance of the engagement of scientists is not to say that only scientists have the expertise to administer or legislate science. To the contrary, scientists generally lack the legislative expertise you have, just as legislators generally lack scientific expertise. Effective scientific legislation therefore benefits enormously from legislators engaging scientists.

In short, any requirements for making data available should carefully consider the complexities, challenges, and potential ramifications. We would be happy to be of any assistance in addressing these concerns, which would require major modifications to the bill.

Sincerely,

A handwritten signature in cursive script that reads "Barry D. Nussbaum".

Barry D. Nussbaum

President, American Statistical Association