COMMENTS OF ROY GAMSE
AT THE UCS-HOSTED PUBLIC HEARING
ON THE EPA SUPPLEMENTAL PROPOSAL
ON STRENGTHENING TRANSPARENCY IN REGULATORY SCIENCE
APRIL 14, 2020

Thank you for creating this opportunity that EPA would not allow. I was Deputy Assistant Administrator of EPA, responsible for overseeing the regulation development process at EPA under two Republican and one Democratic Administrator.

Let me start with what the Supplement does not do. As with the original proposal, it does not provide any reason why this self-regulation is needed. No examples of EPA rules for which underlying data or models are not available to the public and which are faulty as a result. It's just a theory with no supporting evidence, yet one that will be very costly to implement, both in direct costs to implement and very likely in lost benefits to the public.

The costs will be high, though EPA has not provided them, claiming that it is not a major rule costing at least $100 million a year as defined by Presidential Executive Orders. But the Congressional Budget Office estimated that a very similar House of Representatives proposal, The Secret Science Reform Act of 2015, would cost $250 million annually. So where is the economic impact analysis required of every other EPA action over $100MM? That wasn't in the Supplement either.

What is in the Supplement on pages 9 and 10 are two alternatives for dealing with studies for which data or models are not available for independent validation. EPA asks: which do you prefer?

A. tiered access used to reduce the risk of re-identification of private information, or

B. the Agency giving greater weight to studies where the underlying data and models are available than to those for which they are not.

To understand the choice, consider the realities of anonymizing human health data. EPA says it can take a data set of personal health information, obtained in research studies with a promise of confidentiality, and disguise it so the individuals are anonymous. Sounds good, but no longer feasible in these days of "big data analysis."

The International Society of Environmental Epidemiologists' comments submitted on the proposal showed how weak the promise of confidentiality really is when anonymization techniques are used. They showed that:

- In the Harvard Six City Study, most individuals in one of the cities could be identified without name and address information, but just the information needed for independent validation.
- For a Medicare cohort with exposures by ZIP code and the data needed for validation, most of the individuals who died would be identifiable.
- A peer-reviewed study looked at environmental health study in Northern California with data considered by HIPAA to be de-identified, and identified 25% of the participants correctly.
- A study searched a Lexis-Nexis database for stories mentioning hospitalization and identified 43% of the patients without personal identification.

A National Academy of Sciences workshop reached the same conclusion. Attempts to anonymize health data which strip away information that identifies individuals but leave enough for independent replication still allow for identification of the participants.

EPA didn't address the ISEE comments in the Supplement, and the alternatives it proposes don't solve the problem. Chances are that EPA

won't know if the anonymization will indeed protect confidentiality when an ISEE expert or a malicious hacker tries to crack it. So it can use its techniques of anonymization and tiered access to reduce the risk of violating the confidentiality promise, but the promise turns into "best efforts," not a promise at all.

But besides the likelihood of violating privacy promises, future offers of confidentiality to be included in research studies can't be honestly made. If you ask me if my son can participate in a study of lead exposure on intelligence and you offer me "best efforts" at keeping his information secret, my answer is NO and yours would be too. So getting participants in future environmental health studies will get much harder if not impossible.

What about Alternative B, giving epidemiology studies for which personal data is not available a lower priority than other studies? If the epidemiology study is the best study, then it should have the most weight. It is immoral under EPA's governing legislative mandates to not use the best available science, especially due to a rule with no justification.

Look at the track record of health studies used by EPA as the basis for its regulations and ask how many would be given lower or no consideration. Good question. Why hasn't EPA answered it? They would probably be afraid to reveal the answer. Hence, we have a lot of vague statements in the proposal and the Supplement, but no hard data on the real-world impact of the proposals on regulations and health (and as noted above none on the cost of this expensive Major rule).

So my answer to EPA's question, do you prefer alternative A or B is a resounding NEITHER.

I have one more issue to address. On page 26 of Section IV, Availability of Models, EPA says that it is maintaining its proposal that this regulation should apply to data and models evaluated at the time a regulatory action is taken. EPA asks if the rule should apply only to data and models generated after the effective date of the rulemaking.

There is an enormously important question involved. Suspicious minds have suggested that EPA intends for this rule to apply to situations where EPA is required by its legislation to review and update certain regulations, i.e., a stealth attack on existing regulations up for renewal. For example, ambient air quality standards must be revisited every five years. There are air quality standards for which the underlying data cannot be made available because of privacy promises and, in some cases, the data have not been kept. Would EPA give such studies low priority, even though they have been the basis of existing regulations for years? Must EPA really fund new studies just to get a high-priority study to use in updating the old standard? It would be preferable for EPA to clarify that studies completed before this rulemaking are exempt. But this is another case where EPA offers a choice between two horrible rules. The right answer is again a resounding NEITHER.

EPA has not shown any justification of the need for this rule. No examples of what problem is being fixed. No examples of what studies would be downgraded and the effect of doing so. No examples of improperly justified rules. No costs despite the Executive Orders' requirements for them. It has incurred the derision of almost every reputable health and science organization. EPA should stop wasting its and our time on this unnecessary rule.