Comment on EPA Proposed Supplement to "Strengthening Transparency in Regulatory Science" Rule
(In reference to Docket ID No. EPA-HQ-OA-2018-0259)

We are writing on behalf of Concerned Scientists @ IU, a grass-roots, non-partisan community organization consisting of over 1200 members—scientists, students, and supporters of science—from the south-central Indiana region. While many of our members are faculty, students or staff at Indiana University, our organization does not officially represent the University. Concerned Scientists @ IU is dedicated to strengthening the essential role of science in public policy and evidence-based decision making.

We have previously provided a detailed comment (1k2-94fs-ga2c) on the original EPA proposal "Strengthening Transparency in Regulatory Science". Here, we expand on our earlier comments to address in particular the supplement to that proposed rulemaking that EPA now proposes. We continue to **strongly oppose** the proposed rulemaking with the newly proposed supplements because we judge it most likely to *decrease* rather than increase transparency in the use of scientific research in EPA policy making. Our specific objections, on which we expand and amplify in the following paragraphs, center on the following points:

- As in the original proposal, the EPA still fails to make a convincing case that the proposed rulemaking is needed to meet EPA's statutory responsibility to protect public health and the environment. As comments are requested at a time when the ongoing COVID-19 pandemic demonstrates vividly the urgent need to base public health policy on rigorous scientific research, it is unimaginable in this context to consider a proposal to eliminate some such research from consideration.
- While narrowing a few selected definitions, the proposed supplement considerably broadens the scope of the original proposal to now cover not only "pivotal regulatory science," but also "pivotal science" and "influential scientific information." The still vague definitions of these broadened terms creates an even broader opening than the original proposal for political interference in the scientific basis for EPA policies.
- The proposed modification leaves intact the originally proposed authority of the Administrator to grant exemptions from the proposed rule as he/she chooses. If the rule cannot be applied uniformly, then exemptions granted by a political appointee will lead to a widespread loss of public trust in the policy-making process, and thus to a decrease in transparency.
- The proposed supplement justifies its narrow focus on the public availability of raw data by reducing the essential requirement that the research used as a basis for regulatory policy be "reproducible," to stress solely "reanalyzability." This narrowing of focus misses the potential implication that exclusion of research judged by EPA to not be reanalyzable may in fact remove replicated research results from consideration.
- The EPA's stated goal of ensuring greater public availability of the raw data that underlies regulatory decisions is laudable in principle, but in practice best left to scientific peer review and scientific journals, which have strong interest in fostering reproducibility of scientific results. It is not clear that this goal falls within the statutory responsibilities of the EPA.

Lack of a convincing case:

Neither the original proposal nor the proposed supplement provide arguments to demonstrate that the proposed rule will enhance EPA's ability to meet its statutory responsibility to protect public health and the environment. EPA policymaking should be based on the totality of relevant, reproducible scientific

research results. The most detailed studies of health impacts of air and water quality often involve analyses of health records for large numbers of specific individuals, in order to allow for optimal control for other individual health issues that might otherwise cloud conclusions about the specific impact being studied. Such health information is protected from exposure by law. Thus, the need to consider eliminating some research because raw data are not publicly available – sometimes even in a tiered manner – can work against the goal to protect public health.

However, one positive change in the proposed supplement now offers the EPA an opportunity to produce historical research on its own decision-making that could be used to bolster its argument. Specifically, the proposed supplement includes the following: "*If the proposed or alternative approach were finalized, EPA would consider the availability of underlying data and models only for studies that are potentially pivotal to EPA's significant regulatory decisions or influential scientific information that are developed in the future*." Since past regulatory policy adoptions would thus be exempted from the new rule, we feel that EPA should be required to carry out and present an analysis of the impacts this new rule would have had on past regulatory policy, and to demonstrate, in the light of research done since the policy adoption, that those impacts would have *enhanced* the protection of public health and the environment.

Much of the original instigation for industry and political concerns about "transparency" in regulatory science arose from the 1993 Harvard Six Cities Study[1], which analyzed and kept confidential personal health data for 8,111 participating Americans, in order to judge the effects of fine particulate matter ($PM_{2.5}$) concentrations on mortality rates across six American cities. That research was influential in leading to Congressional passage of the 1997 revisions to the National Ambient Air Quality Standards (NAAQS). Would the Six Cities study have been ignored or given little weight as relevant research if the new proposal were applied retroactively? If so, what impact on public health would such elimination have produced?

Numerous post-1997 reanalyses of the Six Studies results and independent studies have since validated the conclusions of the original work judged by some to lack "transparency." For example, a 2004 reanalysis[2] requested by the EPA, industry and non-governmental organizations was "*able to reproduce virtually all of the original numerical results, including the 26 percent difference in overall mortality between the most polluted city (Steubenville, Ohio) and the least polluted city (Portage, Wis.)… Attempts to reproduce the original air-pollution data, for which intact records no longer existed, resulted in some notable discrepancies in the total levels of suspended particulates and the sulfur dioxide levels. However, the discrepancies noted during the audit were not of major epidemiologic importance and did not substantively alter the original risk estimates associated with particulate air pollution nor the main conclusions that were reached."*

More recently, Correia *et al*. analyzed[3] publicly accessible data from 545 U.S. counties that reported annual ambient fine PM levels in 2000 and 2007, along with county-specific life expectancies and other data used to unravel effects of potentially confounding variables, such as socioeconomic status, smoking prevalence and demographic characteristics. Among these 545 counties, they found a strong correlation of reduced life expectancy with increased fine PM levels, and even the small changes in fine PM levels from 2000 to 2007, resulting from attempts to address EPA's revised 1997 standards, caused clearly discernible incremental changes in life expectancy.

Most recently, a new Harvard study[4] has analyzed publicly available data from more than 3000 U.S. counties through April 22, 2020 and "*investigated whether long-term average exposure to fine*

particulate matter ($PM_{2.5}$) is associated with an increased risk of COVID-19 death in the United States." It is, after all, plausible that air pollution that affects respiratory function might also affect susceptibility to the most severe impacts of a respiratory virus. The authors adjusted the data "*by 20 potential confounding factors including population size, age distribution, population density, time since the beginning of the outbreak, time since state's issuance of stay-at-home order, hospital beds, number of individuals tested, weather, and socioeconomic and behavioral variables such as obesity and smoking*." Their conclusions: "*an increase of only 1 $\mu g/m^3$ in $PM_{2.5}$ is associated with an 8% increase in the COVID-19 death rate (95% confidence interval [CI]: 2%, 15%). The results were statistically significant and robust to secondary and sensitivity analyses… Despite the inherent limitations of the ecological study design, our results underscore the importance of continuing to enforce existing air pollution regulations to protect human health both during and after the COVID-19 crisis*."

The recent Harvard study is too new to have been peer-reviewed yet, and is currently available as a preprint. However, its conclusions "*are consistent with previous findings that air pollution exposure increases severe outcomes during infectious disease outbreaks*." Those previous findings have been reviewed in reference 5, and include, for example, studies linking air pollution levels to respiratory disease around a Utah steel mill[6] and to SARS fatality rates in China.[7]

Both the Correia and recent Harvard studies relied on publicly available raw data sorted by county. This leaves such analyses open to question regarding the robustness of adjustments to control for potentially confounding variables, when this is done county-wide. The adjustments and matching of subjects with similar confounding variables but different $PM_{2.5}$ exposure would be more reliable, but the data would be less publicly available, if done for a large sample of health records for individuals. And yet, the proposed EPA rule would more likely eliminate the latter type of study from consideration in policy decisions. We see little in the public record concerning the fine PM standards to suggest that the proposed rule, had it been applied retroactively, would have improved public health, and much to suggest it would have caused a significant deterioration. We expect there are numerous other past EPA regulatory policies to subject to similar retroactive analysis. The burden of proof that the present proposal is needed rests with EPA, and they have not met that burden yet.

Broadened scope of the proposed supplement:

The original proposal focused on the public availability of dose response data and models. The supplement represents a very significant, but inadequately defined, broadening of that scope. As the new proposal states: "*Some, but not the only, examples of information that would be considered to be data and models, in addition to dose-response data and dose-response models, include environmental fate studies, bioaccumulation data, water-solubility studies, environmental fate models, engineering models, data on environmental releases, exposure estimates, quantitative structure activity relationship data, and environmental studies*." The exact scope of the supplement is ambiguous; it is meant to include "*data and models underlying pivotal regulatory science and pivotal science which support significant regulatory decisions and influential scientific information*."

The vagueness of this description gives an EPA Administrator broad license to decide what policy decisions meet the criteria of "pivotal," "significant" and "influential," and therefore, when and how to apply the proposed rule. We judge such license to very likely lead to politically driven suppression of influential research on selected policy areas of particular concern to the administration in power. This opening for political interference in the scientific basis for policy making is only exacerbated by our next concern.

Potential political abuse of exemptions:

The proposed supplement attempts to clarify the only classes of data and models that EPA will be allowed to consider in "pivotal" policy decisions: "*This includes studies with data and models that are publicly available as well as studies with restricted data and models (i.e., those that include confidential business information (CBI), proprietary data, or Personally Identifiable Information (PII) that cannot be sufficiently de-identified to protect the data subjects) if there is tiered access to these data and models in a manner sufficient for independent validation*." However, it reiterates the provision in the original proposal that allows "*the Administrator to grant exemptions from the rule on a case-by-case basis if he or she determines that compliance is impracticable because it is not feasible to ensure that data and models underlying pivotal regulatory science are publicly available in a manner that is consistent with law and protects privacy and confidentiality*."

The granting of such exemptions by a political appointee is likely to make the scientific research basis for EPA policies much *less* transparent than it is currently. For example, the current administration might grant exemptions preferentially to industry-fueled research for which public availability of data and models is compromised by confidential business information. But a future administration might favor exemptions for academic research involving protected PII. A likely result is that the inclusion or exclusion of some research from consideration will become a political, rather than a scientific, issue, to the detriment of public trust, health and the environment.

Narrow focus on reanalyzability:

Public policies should be based on reproducible scientific results. The sole focus in the present proposal on public availability of raw or nearly raw data is coupled to the proposal's narrow definition of "independent validation" as requiring "*that independent analysis of the original or supporting data using identical methods would generate similar analytic results, subject to an acceptable degree of imprecision or error*." But what if independent teams have collected and analyzed independent data, leading them to consistent conclusions? Suppose, as has been the case for the fine PM studies discussed above, that some of these studies have used publicly available county-wide data while others have used protected individual health records. In what way does elimination of the research based on protected data enhance the scientific basis for policy making, when the two approaches have complementary advantages and disadvantages? The proposed rule would seem to reduce reproducible science in such a case to the application of a single approach about which legitimate control questions may be raised.

Our group includes many professional research scientists who have collectively carried out peer review of many thousands of research papers. We are generally able to effectively judge the quality of research by evaluating the completeness and care taken in describing the methods used and conditions imposed on data collection, sample selection, controls exercised, model assumptions and codes used in data analysis, internal consistency of the data, uncertainty analysis, and consistency of the conclusions drawn with the data presented and with earlier relevant results. Only in cases where there appear to be inconsistencies, either internally or with earlier results, would a reviewer insist on trying to reproduce the same analysis of the same data as reported in the paper. The ultimate judgment about the reliability of the results rests with their replication by independent teams using independent data and analysis methods. Why, then, is EPA promoting to such prominence in policy-making the narrow requirement of reanalyzing the same data with identical methods?

Improving public access to data:

Among the topics on which the proposal requests comments is the following: "*EPA is interested in comments about how to provide sufficient incentives and support to researchers to increase access to the data that may be used as pivotal regulatory science or pivotal science*." We believe that the most effective way to do this is to support the role of peer review and scientific journals in their ongoing efforts to enhance the replicability of scientific, and particularly of clinical, research results. Scientific journals have been updating their publication policies to facilitate much more detailed reporting and data archiving,[8] to allow more complete descriptions in published papers of measurement conditions, subject demographics, analysis procedures, assumptions, and computer codes used, as well as to make raw data publicly available, so far as possible in light of privacy and intellectual property concerns.

Substituting the judgment of a political appointee, often with limited background in science, for the professional judgments of the scientific community and journals is not an effective way to increase either public access or public trust in the scientific basis for regulatory decisions. The EPA can, of course, provide incentives to researchers via the level of funding they provide for some of the relevant research. However, the arbitrary exclusion of such EPA-funded research experts from service on EPA scientific advisory boards does tend to compromise the history of good will and professional collaboration the Agency has developed with the scientific community over many decades.

Summary:

We strongly oppose both the originally proposed rule and the newly proposed supplement to it, because we judge the likely outcomes of imposing such a rule to be: (1) political interference in the choice of scientific underpinnings for regulatory policy; (2) *weakened* transparency in regulatory science; and (3) weakening of EPA's statutory commitment to protecting the health and welfare of the American public. We feel that the Agency has done an inadequate job of justifying the need for this proposal.

References:

1. Dockery DW, Pope CA III, Xu X, *et al. An association between air pollution and mortality in six U.S. cities.* N Engl J Med 1993;**329**:1753-1759

2. D. Krewski, *et al., Validation of the Harvard Six Cities Study of Particulate Air Pollution and Mortality,* N Engl J Med 2004; **350**:198-199

3. A.W. Correia, *et al., Effect of Air Pollution Control on Life Expectancy in the United States*, Epidemiology **24**, 23 (2013).

4. Xiao Wu, *et al., Exposure to Air Pollution and COVID-19 Mortality in the United States,* https://www.medrxiv.org/content/10.1101/2020.04.05.20054502v2

5. Ciencewicki J, Jaspers I. *Air pollution and respiratory viral infection*. Inhal Toxicol 2007;**19**(14):1135-46. doi: 10.1080/08958370701665434

6. Pope CA, 3rd. *Respiratory disease associated with community air pollution and a steel mill, Utah Valley*. Am J Public Health 1989;**79**(5):623-8. doi: 10.2105/ajph.79.5.623

7. Cui Y, Zhang ZF, Froines J, *et al*. *Air pollution and case fatality of SARS in the People's Republic of China: an ecologic study*. Environ Health 2003;**2**(1):15. doi: 10.1186/1476-069X-2-15

8. J. Berg, P. Campbell, V. Kiermer, N. Raikhel and D. Sweet, *Joint Statement on EPA Proposed Rule and Public Availability of Data*, Science, Vol. **360**, Issue 6388, eaau0116 (May 4, 2018) [DOI: 10.1126/science.aau0116]