**Science Advisory Board (SAB) Science and Transparency Work Group Questions for EPA Concerning the Proposed Rule *Strengthening Transparency in Regulatory Science***

**May 15, 2019**

The following questions on EPA's proposed rule *Strengthening Transparency in Regulatory Science* were developed by the SAB Science and Transparency Work Group to help EPA and the SAB prepare for discussion of the proposed rule at the SAB meeting held on June 5-6, 2019.

**Science Advisory Board (SAB) Science and Transparency Work Group Questions for EPA Concerning the Proposed Rule *Strengthening Transparency in Regulatory Science***

**May 15, 2019**

## Definition of Data

1. For the purposes of this rule, what is the definition of the "data" underlying a peer-reviewed study? In particular, would it suffice that researchers make available the data on which they performed the bulk of their calculations (the analysis dataset), which typically follows some initial preprocessing or aggregation, or does EPA expect full raw data down to the level of individual measurements, including data directly used to carry out the reported statistical analysis and model development? What level of detail would be provided (e.g., if subjects are removed from the study, would these subjects be identified and reasons given)?

2. How does the role of QA/QC methodology affect the choice of "data" in the stages of aggregation that would be released?

## Validation of Studies

3. How does EPA define "replication," "validation," and "publicly available" for the purposes of this rule?
   - Does "replication" consist of anything other than verifying that applying the same calculations to the same data yields the same results that have been published?
   - Does "validation consist of more than verification of calculations?

4. Given that there are multiple ways to assess validity of epidemiological studies (some of which do not require public access to all data and methods), what does EPA consider to be adequate validation of a study?
   - Would the answer differ if it referenced toxicological studies or environmental characterization studies?
   - Does "validation" encompass validation of interpretations as well as validation of calculations?
   - Should "validation" include testing whether conclusions are robust to changes in methods of analysis?
   - Should methods to allow for third party validation be developed for studies that cannot share data?

## Data Handling

5. Who or what body will bear the data handling costs associated with implementing this regulation? This includes:

- Initial processing and documentation of data prior to public release (including stripping away or suppressing identifying information or confidential business information where necessary);
- Maintenance and administration of data sets so they are "publicly available" (this includes preparing data, beyond what is currently required by a diversity of funding agencies for posting on a public server, oversight of "limited access" data if appropriate, and updates when they are required).

6. How does EPA propose to handle historical data sets that were created long before any of the new rules were put into place, and for which it is not possible to retrospectively apply the proposed procedures?

7. How long do the data sets need to be maintained and publicly available?

**Criteria for Exceptions**

8. What specific criteria would constitute grounds for an exception to the stipulation that data upon which regulation is based be made public?

**Collaboration with other Federal Agencies**

9. The proposed rule contemplates that the "EPA should collaborate with other federal agencies to identify strategies to protect confidential and private information in any circumstance in which it is making information publicly available."
   - What kind of collaboration is contemplated and what is the expected process for collaboration with other agencies?
   - Which other agencies should be involved? Many researchers have worked with data from either the Centers for Disease Control or the Census Bureau, so it would seem logical to include those two agencies at least.

**Consultation with the Science Advisory Board**

10. EPA Administrator Wheeler's April 19, 2019 letter to the Chair of the EPA Science Advisory Board states that "The EPA would benefit from an SAB consultation on existing mechanisms for secure access to confidential business information and personally identifiable information as discussed in the proposal."
    - What form of consultation does EPA envision?
    - What does "existing mechanisms" mean to EPA?