

Additional questions for EPA from SAB members on Science and Transparency Rule issues of secure access to confidential business information (CBI) and personally identifying information (PII).

June 27, 2019

The following additional questions on the issues of secure access to confidential information and personally identifying information were submitted to EPA by members of the Science Advisory Board after the SAB meeting held on June 5-6, 2019.

Additional questions for EPA from SAB members on Science and Transparency Rule issues of secure access to confidential business information (CBI) and personally identifying information (PII).

Questions from Dr. Deborah Bennett

1. Does the EPA have a plan to consider ways to incorporate many of the large international studies that are unlikely to share data with the U.S.?
2. If a study was done in a restrictive data environment, such as a study using census data, and others can apply to replicate the study in that same environment, does the EPA plan to consider that as publicly available?
3. Are there plans for including third party evaluation of studies as an alternative to making data publicly available?
4. Are their plans for studies that must meet HIPPA requirements? In other words, would requirements either be excluded for these or would an alternative process such as third party evaluation be considered?

Questions from Dr. Alison Cullen

1. When are data considered public - specifically, does public mean that every original datapoint is public? Or are approaches in which data are made public but in a clustered or matched form considered to be public for the purposes of this proposed rule? In this approach, subgroups which share characteristics in common would be aggregated and data about these subgroups would be made public. What cost would be associated with this sort of approach?
2. How many regulations are currently based on past studies which are not fully publicly available? How many of these total studies are not fully public due to confidential business information and how many are not fully public due to reasons related to personal identification issues? What are the relative magnitudes of these two barriers to publicizing information currently supporting regulation?
3. If the answer to #2 is not known or not easily established, what is the anticipated cost and effort associated with identifying which specific regulations are based on not fully public data (and which specific datasets underlie these, as well as the reason for their non-public nature)? Who would be responsible for that cost? What process would be used to be sure that there is a comprehensive understanding of the extent of the issue of not fully public data across regulations (rather than a focus on a small set of singled out regulations and not fully public datasets)?

4. In light of the above questions, if this regulation moves forward, what would be the cost of making datasets fully publicly available, what are the characteristics, complexity and sample sizes of these datasets? What party or body would bear those costs?