COMMENTS BY THE TEXAS COMMISSION ON ENVIRONMENTAL QUALITY REGARDING STRENGTHENING TRANSPARENCY IN REGULATORY SCIENCE PROPOSED RULE

EPA DOCKET ID NO. EPA-HQ-OA-2018-0259

I. Summary of Proposed Action

On April 30, 2018, the United States Environmental Protection Agency (EPA) published a proposed rule in the *Federal Register* (83 FR 18768) titled *Strengthening Transparency in Regulatory Science*. The EPA provided a 30-day public comment period, ending on May 30, 2018. On May 25, 2018, the EPA extended the public comment period to August 16, 2018 (83 FR 24255). This proposed rule (hereafter referred to as the proposed data transparency rule) would establish an EPA policy that would only allow scientific data that is publicly available and independently verifiable to be used as the basis for significant regulations. The EPA specifically requested comment on numerous details related to the proposed data transparency rule's scope and implementation.

II. Comments

A. General Comments

The Texas Commission on Environmental Quality (TCEQ) appreciates the EPA's intention to provide greater transparency to the broader community of scientists, regulators, regulated entities, and interested members of the public who rely on and are impacted by the EPA's regulations. We encourage the EPA to interact with experts in each of these areas throughout the development process to ensure that all the opportunities and challenges presented with this proposed data transparency rule are fully realized and considered.

The TCEQ appreciates the EPA's commitment to having a strong scientific foundation for environmental regulations. We agree that many of the EPA's environmental regulations impact the daily lives of Americans and the importance of maintaining their trust through our shared scientific integrity cannot be overstated. The proposed rule emphasizes the need for major environmental regulations to be comprehensive, unbiased, and transparent. Because this is the first of its kind both within the EPA and among federal agencies, we hope that the EPA considers this proposed data transparency rule to be one step toward a longer conversation about these shared objectives.

The proposed data transparency rule language was silent on several important technical considerations, presumably to allow the public an opportunity to shape the rule. However, the lack of specific policy design has led to confusion among experts and particularly the media about the real consequences of this proposed rule. Numerous news articles have decried the expected significant loss of scientific studies that this rule may cause to be excluded from future analyses. In contrast to this view, the proposed data transparency rule could be used to continue to drive scientific investigation by allowing researchers from different backgrounds to re-analyze important datasets. Critical review, reanalysis, and replicable results are cornerstones of science. Further, the global scientific community has already been discussing the importance of the availability of de-identified data for reanalysis since at least 2010 (IOM 2012). It is also consistent with the EPA's Enterprise Architecture and Quality System for Environmental Data

and Technology that is committed to improving data quality and promotion of data sharing. Several journals (e.g., Public Library of Science (PLoS), the Annals of Internal Medicine) already have a publication condition that requires authors to make their data available upon request. In addition, many repositories already exist for sharing certain data, including clinical trial data. Just a few examples include the Yale University Open Data Access (YODA) project, the Harvard Library Dataverse Project, Dryad, and the Health and Medical Care Archive. Although there may be challenges to gathering, storing, and releasing de-identified data, a thoughtfully implemented data transparency rule could build upon the scientific community's growing trend and compel current and future generations of scientists to share their data, leading to stronger scientific understanding and more meaningful regulations.

The potential opportunities and challenges of the proposed data transparency rule are numerous and should be vetted through existing experts. Rather than collect and respond to numerous disjointed comments, it would benefit the EPA and the rule to engage the expertise of the many organizations that have vast experience with collecting, storing, and sharing confidential information. As such, the TCEQ strongly encourages the EPA to convene a work group or review panel of experts to help in guiding the agency on several of the important details that are needed to make this proposed data transparency rule succeed. The work group should comprise diverse perspectives, including members from other federal agencies (e.g. the United States Department of Health and Human Services; the Health Services Advisory Group; the National Academies of Sciences, Engineering, and Medicine; the National Institutes of Health) and institutional review boards who would best be able to discuss the most recent and relevant methods for collecting and sharing health data from human studies. The work group should also examine economic and environmental modeling data, the extent to which data should be replicable, necessary privacy restrictions, exception criteria, and the potential for inadvertent bias due to the proposed data transparency policy. The EPA should then provide a more explicit proposed data transparency rule for public inspection and comment. Taking these steps will lead to a more thoughtful and meaningful policy that promotes and is led by scientific research while providing necessary regulatory transparency.

B. Specific Comments

Governance

The TCEQ strongly encourages the EPA to give governing authority for granting exceptions to the proposed data transparency rule, as well as the oversight of raw data collection, storage, and access to an external entity or entities to ensure independence and objectivity.

The proposed data transparency rule solicits public comment on key implementation details, such as criteria for granting exceptions to the rule, how to ensure an appropriate balance of regulatory transparency and data protection, and a data sharing platform. Consistently and thoughtfully granting exceptions will be one of the most crucial steps in ensuring the success of the data transparency rule, and more importantly the ongoing success of using science to inform policy. The true tone of this proposed rule will be set through exception determinations that either exclude certain evidence or justify using non-accessible data. Exceptions will also ensure that regulatory action continues to take place in the event that the EPA has a statutory requirement to make a decision using data that are not available (e.g., an older study for which the original dataset no longer exists) or cannot be released (e.g., confidential business information). Because the EPA has a vested interest in the outcome of such a decision, the TCEQ recommends that exception decisions rest with an external, third-party entity, such as the EPA's Science Advisory Board or a new cross-disciplinary board. As with many ethical determinations, organizational independence provides greater trust that decisions were reasonable, objective,

and unbiased. The governing entity could keep abreast of upcoming regulations and determine whether the pivotal scientific methods, data, or models used in the regulation's assessment document(s) would be subject to this proposed data transparency rule. Over time, the entity could develop a list of pivotal scientific articles, methods, data, or models that could then be available to the public on an open access website. Documenting pivotal works would help ensure consistent application of the rule and guide researchers in further framing and building their own research agendas. To be successful, the governing entity should be given limited authority to enforce its exclusion determinations. The balance between data privacy and regulatory transparency would best be respected through both this governing body and the prioritization of regulatory actions that fall under this rule.

Depending upon the identity and role of this governing entity, the same body or a separate third-party entity should oversee the storage of the data and grant access to qualified researchers. Other organizations already have infrastructure and policies in place to handle sensitive information, and it makes sense to use these existing capabilities, rather than attempting to re-create them at the EPA. In addition to the convenience, though, the independence of the entity that would grant access to the data also ensures that the EPA is not put in the paradox of collecting data to ensure data transparency and then possibly having to reject requested access to that data. More specific comments related to data accessibility are provided in the last section of this document.

Scope and Timing

The EPA should focus this proposed rule on a narrow set of actions that includes "major" or "economically significant" regulations and should apply data transparency rule requirements to documents that inform a proposed rule so that the public has adequate time to evaluate the data. This will better ensure timely regulatory action, as well as balance the need for confidentiality with regulatory transparency.

Although transparency is important throughout all regulatory actions, increasing levels of transparency can concomitantly complicate the review process, particularly when the data include sensitive personal information or confidential business information. The additional scientific and legal review, considerations, analyses, and approvals necessary to validate pivotal studies will likely lengthen the rule development and external review process.

The challenge, then, is to balance the need for regulatory transparency with preservation of sensitive information *and* the need for timely regulatory action. This can be accomplished in part through restricting the scope of the proposed data transparency rule to those actions that have the greatest impact on public health and regulated entities. Specifically, those rules that are considered "major" under the Congressional Review Act or "economically significant" under Executive Order 12866. Regulatory actions outside of this scope (e.g., site-specific permitting actions) may still impact the public, though the impact is much less than that of major or economically significant rules.

In order to have meaningful public involvement in the development of major rules, the EPA should apply the requirements from this data transparency rule as early in the process as is scientifically practical. Waiting until a final rule (or even a proposed rule, in the case of some rule-makings) to notify the public of pivotal research would leave no time to determine whether data could be accessed or for the public to review or comment on the revised analysis, should the EPA exclude a key study or be granted an exception from this data transparency rule. Further, if subsequent post-rule validation attempts prove unsuccessful, the EPA would then be put in the difficult position of repealing a rule that would likely have already cost state governments and regulated entities capital and resources to implement. To ensure that the goal of greater

transparency does not stymie the EPA's already extensive scientific review of available literature, the EPA should continue its existing evaluation process of surveying all possible materials first and providing its technical support documents or assessment documents for public inspection. Once available studies have been reviewed, pivotal scientific evidence (e.g., articles, models, etc.) can be identified and either retrieved or granted an exception through a standardized review process conducted by the third-party governing authority. A subsequent draft of the technical support or assessment document should then be provided if the exclusion of pivotal evidence required additional review or reliance on other, formerly non-pivotal, scientific evidence.

The proposed data transparency rule should also apply to scientifically or technically novel guidance, and the resulting guidance document should be made available for public comment.

If the EPA chooses to issue guidance that includes new scientific or technically novel approaches and intends to impose compliance on state, tribal, and local governments and grantees tasked with implementing environmental quality rules, the guidance document should be subject to this proposed data transparency rule and should be made available for public review and comment. Allowing this opportunity for public involvement safeguards the larger goal of regulatory transparency, as operational guidelines can often be just as impactful as official regulatory actions.

The EPA should consider pivotal scientific research as rules are formally reviewed, rather than retrospectively.

The EPA requested comment on whether the proposed data transparency rule should be applied retroactively to rules that have already been promulgated. Retroactive application of this proposed rule to existing rules would be infeasible due to the volume of significant environmental rules already in place and the time and expertise it would take to reanalyze the technical and legal aspects of these rules while maintaining oversight of ongoing agency business. Fortunately, for the most part, retroactive application is likely unnecessary. Many of the EPA's significant rules have already been implemented and are on a statutorily-mandated review cycle. Once the EPA solidifies the requirements of this proposed data transparency rule, they should commit to applying the new transparency standards to pivotal scientific studies, both old and new, that justify agency actions on rules as they are formally reviewed on their existing cycle. In a separate analysis, the EPA should evaluate whether any rules that are not on a review cycle should be re-evaluated through the lens of this data transparency rule and the EPA should receive feedback on their determinations from the public and any applicable review boards or entities.

The EPA should consider using an upcoming major or significant rule-making as a test case to help solidify how the intent of the data transparency rule could be thoughtfully implemented before phasing in the final data transparency rule.

Because the proposed data transparency rule offers limited implementation details, it is difficult to provide meaningful comment on the potential intentional and unintentional consequences of its enactment. Rather than promulgate an untested final rule that could potentially impact all agency actions, the EPA should consider choosing a major upcoming rule-making as a test case. Ideally, the test case would help sketch out more specific guidelines for data transparency rule implementation that could then be vetted with an external, cross-disciplinary work group and provided for public review and comment. As stated above, once more solidified guidance is available the EPA should phase in the final data transparency rule by ensuring that data transparency guidelines are followed during the formal review process for new and existing major rules.

Dose-Response Models

The TCEQ applauds the proposed rule language relating to consideration of dose-response models and uncertainty, rather than relying on default assumptions.

The TCEQ strongly supports the EPA considering multiple dose-response models based on information such as biological plausibility, mode of action, mechanism of action, etc. that is relevant to the selection of the most scientifically-appropriate model(s). Biologically-based models (e.g., the formaldehyde CIIT model) should be explicitly included for consideration in addition to those listed in the proposal (i.e., linear, threshold, and U-shaped, J-shaped, bell-shaped). Deviations from the use of default models should be evaluated on a case-by-case basis and have adequate scientific justification for use of an alternative model better supported by the chemical-specific data. Adequate scientific justification to include an alternative model in an assessment should not be an ever-changing and insurmountable standard, but rather should be based on an objective evaluation of the scientific weight-of-evidence of relevant and robust data. This approach is consistent with regulatory agencies encouraging new research to generate data to replace conservative defaults. Strong, science-based regulations should always be guided and driven by the actual data at hand as opposed to mere defaults.

Definitions

Where possible, the EPA should be explicit about important definitions, including "publicly available," "pivotal regulatory science," "dose response data and models," and "validation."

The lack of explicit definitions in the proposed data transparency rule has caused confusion and uncertainty among experts outside the EPA. The broadest, and likely unintentional, interpretation of several key terms used in this proposed rule could completely stall or alter the existing state of the EPA's scientific and regulatory process. A hasty and incorrect assumption made by many has been that this concern makes the entire proposed rule faulty and unworkable. Some of this confusion will be cleared with more detailed implementation information; however, the EPA should also provide more explicit definitions of key terms wherever possible.

Perhaps one of the more frequently used terms that has caused the most confusion is "publicly available." Outside the context of this proposed rule, "publicly available" generally means that information is available without restriction to anyone who wants it. While this level of accessibility is acceptable for most information, it becomes problematic for sensitive personal information, personal health information, or confidential business information. In order to meet existing privacy and confidentiality laws that protect this type of information, increased accessibility tends to come with reduced refinement (i.e., less detailed information) in the data. For many health studies, data could become so unrefined as confidential data are removed from the dataset in order to become publicly available that results would be incomparable to those in the original study, and therefore not replicable or verifiable. The Institute of Medicine (IOM) defines and provides examples of three varying forms of data sharing that the EPA should consider in better framing their intended accessibility of health and study data (IOM 2015). The TCEQ recommends that the EPA strongly consider implementing a formal request and approval process for qualified researchers to obtain access to data that is protected by confidentiality laws, rather than aiming for all data to have unrestricted access. This process would still allow identifiable or fairly refined de-identified data to be re-evaluated by experts external to the EPA or the original study authors, while protecting the privacy of the study participants or businesses who created the model or dataset.

Other definitions that would be helpful in framing the proposed rule include "pivotal regulatory science," "dose response data and models," and "validation." Although "pivotal regulatory science" is likely intended to mean just the key study or studies underlying a regulatory analysis, it could be argued that supporting studies (i.e., those that help to build the weight of evidence or describe the mode of action) are also pivotal to the analysis. In addition, it is unclear what the term "dose response data and models" includes. This term could mean anything from summary data to specific calculations and code, which matters for the potential sensitivity of the information. Finally, the EPA should define what it means to "validate" study results. This could be interpreted to mean that the study must be repeated (i.e., a new group of human subjects are exposed in a controlled experiment or are assessed in an observational setting) or simply that the statistics and models are recalculated from the existing data. Further, the EPA should determine what level of validation is necessary for pivotal regulatory science so that it can be consistently applied.

The TCEQ applauds the EPA's intention to make regulatory science more transparent.

References

United States Department of Health and Human Services (HHS). 2012. Guidance regarding methods for deidentification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Washington, DC: HHS.

Institute of Medicine (IOM). 2015. Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk. Committee on Strategies for Responsible Sharing of Clinical Trial Data; Board on Health Sciences Policy; Institute of Medicine. Washington (DC): National Academies Press (US); 2015 Apr 20.