Redline Revisions --- Revised to address Interagency Comments received under EO 12866 review

6560-50-P

ENVIRONMENTAL PROTECTION AGENCY

40 CFR Part 30

[EPA-HQ-OA-2018-0259; FRL-10004-72-ORD]

RIN 2080-AA14

Strengthening Transparency in Regulatory Science

AGENCY: Environmental Protection Agency (EPA).

ACTION: Supplemental notice of proposed rulemaking.

SUMMARY: This supplemental notice of proposed rulemaking (SNPRM) includes clarifications, modifications and additions to certain provisions published on April 30, 2018. This SNPRM proposes that the scope of the rulemaking apply to influential scientific information as well as significant regulatory decisions. This notice proposes definitions and clarifies that the proposed rulemaking applies to ~~all~~ data and models underlying both pivotal science ~~used to support decision making.~~and pivotal regulatory science. In this SNPRM, EPA is also proposing ~~alternate approaches~~a modified approach to the public availability provisions for data and models that would underly significant regulatory decisions and an alternate approach. Finally, EPA is ~~proposing 5 U.S.C. 301 as sole~~taking comment on whether to use its housekeeping authority independently or in conjunction with appropriate environmental statutory provisions as authority for taking this action.

DATES: Comments must be received on or before [*Insert date 30 days after date of publication in the* ***Federal Register***]

ADDRESSES: You may send comments, identified by Docket ID No. EPA-HQ-OA-

2018-0259, by any of the following methods:

Federal eRulemaking Portal: https://www.regulations.gov/ (our preferred method). Follow the online instructions for submitting comments.

Mail: U.S. Environmental Protection Agency, EPA Docket Center, Office of Research and Development Docket, Mail Code 28221T, 1200 Pennsylvania Avenue NW, Washington, DC 20460.

Hand Delivery / Courier: EPA Docket Center, WJC West Building, Room 3334, 1301 Constitution Avenue, NW, Washington, DC 20004. The Docket Center's hours of operations are 8:30 a.m. – 4:30 p.m., Monday – Friday (except Federal Holidays).

Instructions: All submissions received must include the Docket ID No. for this rulemaking. Comments received may be posted without change to https://www.regulations.gov/, including any personal information provided.

FOR FURTHER INFORMATION CONTACT: Cheryl A. Hawkins, Office of Science Advisor, Policy and Engagement (8104R), Environmental Protection Agency, 1200 Pennsylvania Ave NW, Washington, DC 20460; telephone number: (202) 564-7307; email address: osp_staff@epa.gov***.***

SUPPLEMENTARY INFORMATION:

I. **Executive Summary**

A. Does this action apply to me?

This SNPRM does not regulate any entity outside the Federal Government. Rather, the proposed requirements would modify the EPA's internal procedures regarding the transparency of science underlying regulatory decisions. However, the Agency recognizes that any entity interested in EPA's regulations may be interested in

this proposal. For example, this proposal may be of particular interest to entities that conduct research ~~and other~~or another scientific activity that is likely to be relevant to EPA's regulatory activity.

B. What is the Agency's authority for taking this action?

On April 30, 2018, EPA published the Strengthening Transparency in Regulatory Science Proposed Rulemaking ("2018 proposed rulemaking," Ref. 1). The 2018 proposed rulemaking cites as authority several environmental statutes that EPA administers: the Clean Air Act; the Clean Water Act; the Safe Drinking Water Act; the Resource Conservation and Recovery Act; the Comprehensive Environmental Response, Compensation, and Liability Act; the Federal Insecticide, Fungicide, and Rodenticide Act; the Emergency Planning and Community Right-To-Know Act and the Toxic Substances Control Act. Subsequently, in the Federal Register at 83 FR 24255, May 25, 2018, EPA published a document extending the comment period and announcing a public hearing on the 2018 proposed rulemaking to be held on July 18, 2018 (Ref. 2). That document identified 5 U.S.C. 301 as a source of authority in addition to those statutes cited in the 2018 proposed rulemaking. With respect to the authorities cited in the 2018 proposal, EPA is clarifying that the citation to the Resource Conservation and Recovery Act ("RCRA") section 7009, 42 U.S.C. 6979, should be to RCRA section 8001, 42 U.S.C. 6981; the citation to the Comprehensive Environmental Response, Compensation, and Liability Act ("CERCLA") section 116, 42 U.S.C. 9616, should be to CERCLA section 115, 42 U.S.C. 9615; and including the Clean Water Act section 501, 33 U.S.C. 1361.

EPA is authorized to promulgate this regulation under ~~the Federal Housekeeping~~

~~Statute, 5 U.S.C. 301~~its housekeeping authority. The Federal Housekeeping Statute provides that "[t]he head of an Executive department or military department may prescribe regulations for the government of his department, the conduct of its employees, the distribution and performance of its business, and the custody, use, and preservation of its records, papers, and property."

~~On April 30, 2018, EPA published the Strengthening Transparency in Regulatory Science Proposed Rule ("2018 proposed rule," Ref. 1). The 2018 proposed rule cites several environmental statutes that EPA administers as authority: Clean Air Act; Clean Water Act; Safe Drinking Water Act; Resource Conservation and Recovery Act; Comprehensive Environmental Response, Compensation, and Liability Act; Federal Insecticide, Fungicide, and Rodenticide Act; Emergency Planning and Community Right-To-Know Act and Toxic Substances Control Act. Subsequently, on May 25, 2018, EPA published a notice extending the comment period and announcing a public hearing on the 2018 proposed rule to be held on July 18, 2018 (Ref. 2). That notice identified~~ 5 U.S.C. 301 ~~as a source of authority in addition to those statutes cited in the 2018 proposed rule.~~

~~Section 301 provides appropriate authority for agencies to promulgate regulations that govern internal agency procedures.~~. As the Supreme Court discussed in *Chrysler Corp v. Brown*, the intended purpose of section 301 ~~is~~was to grant ~~a federal agency~~early Executive departments the authority "to ~~regulate its own~~govern internal departmental affairs~~."~~."[1] ~~.~~ As the Supreme Court further notes, section 301 authorizes

[1] *Chrysler Corp. v. Brown*, 441 U.S. 281, 309 (1979).

"what the [Administrative Procedure Act] terms 'rules of agency organization, procedure or practice' as opposed to substantive rules."[2]

EPA is not one of the 15 "Executive Departments" listed at 5 U.S.C. 101. However, EPA gained housekeeping authority through the Reorganization Plan No. 3 of 1970, 84 Stat. 2086 (July 9, 1970). The Reorganization Plan created EPA, established the Administrator as "head of the agency" and transferred functions and authorities of various agencies and Executive departments to EPA.

Section 2(a)(1)-(8) of the Reorganization Plan transferred to EPA functions previously vested in several agencies and executive departments including the Departments of Interior and Agriculture. Section 2(a)(9) also transferred so much of the functions of the transferor officers and agencies "as is incidental to or necessary for the performance by or under the Administrator of the functions transferred."

The Office of Legal Counsel has opined that the Reorganization Plan "convey[s] to the [EPA] Administrator all of the housekeeping authority available to other department heads under section 301" and demonstrates that "Congress has vested the Administrator with the authority to run EPA, to exercise its functions, and to issue regulations incidental to the performance of those functions."[3]

Courts have considered EPA to be an agency with section 301 housekeeping authority. The U.S. Court of Appeals for the Second Circuit, in *EPA v. General Elec. Co.*, 197 F.3d 592, 595 (2d Cir. 1999), found that "the Federal Housekeeping Statute, 5 U.S.C. 301, authorizes government agencies such as the EPA to adopt regulations

[2] *Id.* at 310.

[3] Authority of EPA to Hold Employees Liable for Negligent Loss, Damage, or Destruction of Government Personal Property, 32 O.L.C. 79, 2008 WL 4422366 at *4 (May 28, 2008) ("OLC Opinion")

regarding 'the custody, use, and preservation of [agency] records, papers, and property.'" The Fourth Circuit Court of Appeals, in *Boron Oil Co. v. Downie*, 873 F.2d 67, 69 (4th Cir. 1989), held that the district court exceeded its jurisdiction where it compelled testimony contrary to duly promulgated EPA regulations which EPA argued were authorized by section 301.

EPA's housekeeping authority was established by the Reorganization Plan. As indicated by the case law and the OLC Opinion, it has long been recognized that EPA has been granted full section 301 or equivalent authority. Therefore, EPA has ample authority to promulgate regulations that govern internal agency procedures.

The 2018 proposed rulemaking, as supplemented by this SNPRM and this accompanying preamble, describes how EPA will ~~ensure that~~handle studies when data and models underlying science that is pivotal to EPA's significant regulatory decisions ~~are~~or influential scientific information are or are not publicly available in a manner sufficient for independent validation and analysis. ~~In addition, this supplemental proposal and this accompanying preamble describe how EPA will use pivotal regulatory science and its underlying data and models in developing EPA's significant regulatory decisions.~~ The rule would not regulate the conduct or determine the rights of any entity outside the federal government.[4] Rather, it exclusively pertains to the internal practices of the EPA.

Finally, EPA in the 2018 proposed rulemaking, as supplemented by this SNPRM

[4] *See also United States v. Manafort*, 312 F. Supp. 3d 60, 75 (D.D.C. 2018) (explaining that the Department of Justice "was not at all ambiguous about what it was doing when it promulgated the Special Counsel Regulations [under the authority of 5 U.S.C. 301~~,~~], and it emphasized that it was not creating a substantive rule.").

and this accompanying preamble, does not propose to interpret provisions of~~, nor does it propose to exercise substantive rulemaking authority delegated to it by,~~ a particular statute or statutes that it administers. Instead, in this action, EPA proposes a rule of agency procedure to establish an agency~~-~~ wide approach to ~~ensure that~~handling studies when the data and models underlying EPA's significant regulatory decisions and influential scientific information are publicly available~~.~~ and when those data and models are not publicly available. Therefore, this is a proposed internal rule of agency procedure ~~under EPA's section 301 authority~~.

This internal agency procedure is intended to be consistent with the statutes that EPA administers and EPA plans to implement this procedural rulemaking in accordance with all applicable statutory and regulatory requirements. ~~Therefore~~Indeed, as discussed in this SNPRM, EPA is also proposing options that would allow EPA to consider studies even if the underlying data and models are not publicly available. [See Section IV.] The Agency seeks comment on whether this approach may improve consistency between this proposed rulemaking and certain provisions of those statutes that refer to standards for data availability. Nonetheless, in the event the procedures outlined in the proposed ~~rule~~rulemaking conflict with the statutes that EPA administers, or their implementing regulations, the statutes and regulations will control. Moreover, EPA is considering how to proceed, apart from this supplemental proposal, to establish regulations interpreting provisions of, and/or exercising substantive rulemaking authority delegated to it by programmatic statutes, to include one or more of those statutes cited as authority in the 2018 proposed ~~rule. However, as set forth above, EPA does not intend to rely on those statutes as authority for the 2018 proposed rule as supplemented~~

~~by this supplemental proposal and this accompanying preamble~~rulemaking as clarified in this SNPRM.

Although this is a rule of internal agency procedure and EPA does not propose to interpret provisions of a particular statute or statutes that it administers, EPA is taking comment on whether to use its housekeeping authority independently as authority or in conjunction with the environmental statutory provisions cited as authority in the 2018 proposed rulemaking (as clarified in this SNPRM). The Agency continues to consider whether it is appropriate to rely on its authority in the above-reference environmental statutory provisions (potentially in conjunction with its housekeeping authority). The Agency will consider comments on this issue submitted in response to the 2018 proposal and in response to this SNPRM.

C. What action is the Agency taking?

EPA is issuing this SNPRM to clarify, modify and supplement certain provisions included in the 2018 proposed ~~rule. EPA~~rulemaking in response to some of the public comments received on the 2018 proposed rulemaking (83 FR 18768), as well as to ensure consistency with the April 2019 release of the White House's Office of Management and Budget (OMB) Memorandum to the Heads of Executive Departments and Agencies entitled Implementation of the Information Quality Act (OMB M-19-15, Ref. 3). This memorandum is ~~also proposing 5 U.S.C. 301 as sole authority~~directly relevant to several of the provisions of the 2018 proposed rulemaking because it updates implementation of OMB's 2002 Guidelines for ~~taking this action.~~Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies to, among other things, reflect recent innovations and policies

surrounding information access.

First, EPA is modifying the regulatory text initially proposed in the 2018 proposed rulemaking at 40 CFR 30.3, 30.5, 30.6 and 30.9 so that these provisions would apply to ~~all~~ data and models, not only dose-response data and dose-response models. In addition, EPA is clarifying that the use of the terms "model assumptions" and "models" in the proposed regulatory text at ~~§ 30.6 apply to every assumption used (*e.g.*, body weight) within an individual model.~~ 40 CFR 30.6 apply to the assumptions that drive the model's analytic results. EPA has modified the regulatory text at 40 CFR 30.6 to reflect this clarification. This approach is consistent with OMB M-19-15 (Ref. 3), which highlights the need to characterize the sensitivity of an agency's conclusions to analytic assumptions.

Second, EPA is ~~modifying~~proposing to expand the scope of this rulemaking to apply to influential scientific information[5,6] as well as significant regulatory actions. EPA is proposing to add definitions for "influential scientific information" and "pivotal science" at 40 CFR 30.2 that will pertain to the science underlying influential scientific information, which are not regulatory, and is making conforming changes to proposed 40 CFR 30.3, 30.5, 30.6 and 30.7. EPA is retaining the definition of "pivotal regulatory science" from the 2018 proposed rulemaking regulatory text.

[5] The term "influential scientific information" means scientific information the agency reasonably can determine will have or does have a clear and substantial impact on important public policies or private sector decisions (OMB M-05-03). A "highly influential scientific assessment" is a subset of influential scientific information and refers to "an evaluation of a body of scientific or technical knowledge that typically synthesizes multiple factual inputs, data, models, assumptions, and/or applies best professional judgment to bridge uncertainties in the available information" and that the dissemination of such assessment could have "a potential impact of more than $500 million in any one year on either the public or private sector or that the dissemination is novel, controversial, or precedent-setting, or has significant interagency interest" (OMB M-05-03).

[6] See EPA's Peer Review Agenda at https://cfpub.epa.gov/si/si_public_pr_agenda.cfm

Third, EPA is modifying, deleting and proposing new regulatory text in addition to proposing definitions for "influential scientific information" and "pivotal science" at proposed 40 CFR 30.2. EPA is deleting the first paragraph of the 2018 proposed ~~rule~~rulemaking regulatory text at 40 CFR 30.2 ~~and~~. EPA is deleting the definition of "research data" at 40 CFR 30.2. EPA is proposing definitions for the terms "~~reanalyze~~capable of being substantially reproduced," "data," "independent validation," "~~data," and "~~model," "publicly available," and "reanalyze." These revisions are intended to provide clarity on key terminology used in the regulatory text in the 2018 proposed ~~rule~~rulemaking as well as in this supplemental proposal.

~~Third, in addition to the changes to~~ Fourth, EPA is deleting the 2018 proposed ~~rule~~regulatory text at ~~§ 30.5 described earlier, *i.e.*, broadening its applicability~~40 CFR 30.10. This change is being made to be consistent with the deletion of "research data" in 40 CFR 30.2 because 40 CFR 30.10 would have required EPA to implement this rulemaking to be consistent with the definition of "research data." With the deletion of "research data" from ~~"dose-response data and models" to all data and models~~proposed 40 CFR 30.2, proposed 40 CFR 30.10 is no longer needed.

Fifth, EPA is proposing ~~two additional alternate approaches to~~ a modified version of the regulatory text at 40 CFR 30.5 from that proposed in the 2018 proposed rulemaking. Under this new approach to proposed 40 CFR 30.5~~. The first additional alternate option EPA is proposing would be to use the public availability of the data and models as an important factor in determining whether the agency should utilize certain studies. Whether the underlying data for a study~~, when promulgating significant regulatory decisions or finalizing influential scientific information, the ~~computer code or~~

~~data underlying a model were publicly available would be weighed with the other assessment factors identified in Unit IV.A of this preamble to determine whether the studies based on the data or models could be used as~~ Agency will only use pivotal regulatory science~~. This would apply to all data and models regardless of when they were generated (*i.e.*, when the development of the data set or model has been completed or updated). In addition, EPA is requesting comment on a variation of this option that would apply only to data and models generated after the effective date of the final rule for this rulemaking. Another alternate option to proposed § 30.5 would provide for tiered access to~~ and/or pivotal science if the data and models are available in a manner sufficient for independent validation. This includes studies with data and models that ~~have~~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 ~~anonymized, and require that all other~~ sufficiently de-identified to protect the data subjects) if there is tiered access to these data and models ~~be made~~ in a manner sufficient for independent validation. Tiered access includes the appropriate techniques used to reduce the risk of re-identification and, therefore, mitigate certain disclosure privacy risks associated with providing such access.

As an alternative, EPA is proposing that under proposed 40 CFR 30.5, when promulgating significant regulatory decisions or finalizing influential scientific information, other things being equal, the Agency will give greater consideration to studies where the underlying data and models are available in a manner sufficient for independent validation either because they are publicly available ~~if they are to be used~~

~~as pivotal regulatory science.~~or because they are available through tiered access when the data includes CBI, proprietary data, or PII that cannot be sufficiently de-identified to protect the data subjects. The Agency will identify those studies that are given greater consideration and will provide a short description of why greater consideration was given. As discussed later in the preamble, such approaches to increasing access to data and models can often allow stakeholders to reanalyze the data and models and explore the sensitivity of the conclusions to alternative assumptions while accessing only the data and aspects of the models that they need. This ~~option~~proposal would apply to ~~all~~reviews of data ~~and~~, models, and studies at the time a rule is developed or influential scientific information is finalized, regardless of when the data and models were generated. See Section IV of this preamble for a description of these ~~alternate~~ proposals.

~~Fourth~~Sixth, EPA is modifying 40 CFR 30.9 to describe the factors the Administrator would consider in determining whether to grant an exemption to the proposed public availability requirements for using data and models in significant regulatory decisions and influential scientific information.

~~Finally,~~Seventh, the EPA is proposing the ~~use~~option of ~~5 U.S.C. 301 (the "Federal Housekeeping Statute") as the sole statutory~~ using its housekeeping authority ~~for taking this action. Under this proposal, EPA would no longer be citing the substantive statutes identified in the 2018 proposed rule~~independently as authority for taking this action~~.~~ or in conjunction with the environmental statutory provisions cited as authority in the 2018 proposed rulemaking (as clarified in this supplemental proposal). The Agency continues to consider whether it is appropriate to rely on its authority in the

above-referenced environmental statutory provisions (potentially in conjunction with its housekeeping authority). The Agency will consider comments on this issue submitted in response to the 2018 proposal and in response to this SNPRM. Section 301 authority as transferred to EPA in Reorganization Plan No. 3 of 1970 provides appropriate authority for ~~agencies~~EPA to promulgate regulations that govern internal agency procedures. This action establishes internal agency procedures governing how EPA employees will ~~ensure that~~handle studies when the data and models underlying science that is pivotal to EPA's significant regulatory decisions ~~is~~and/or influential scientific information are or are not publicly available.

The 2018 proposed rulemaking solicited comment on all aspects of the proposed rulemaking. This SNPRM solicits comment only on the changes and additions to the proposed regulatory text discussed in this supplemental document. Comments submitted in response to this supplemental document that address aspects of the 2018 proposed rulemaking that are not addressed, altered, or replaced by this SNPRM will be deemed outside the scope of this supplemental action.

D. Why is the Agency taking this action?

EPA received extensive comment on the 2018 proposed rulemaking regarding the clarity of certain aspects of the proposed rulemaking and the challenges in making all dose-response data and models publicly available. The intent of this supplemental proposal SNPRM is to address certain concerns raised about the clarity of the 2018 proposed rulemaking; to clarify consistency with OMB M-19-15, OMB M-05-03 (Final Information Quality Bulletin for Peer-Review, Ref. 4), and Executive Order 13891 (Ref. 5); to propose an alternate ~~approaches to the scope of the~~ 40 CFR 30.5 provision for

~~public~~ availability of data and models underlying pivotal regulatory science and pivotal science, and to propose relying on 5 U.S.C. 301 ~~as the sole~~independently or in conjunction with the environmental statutory provisions cited as authority ~~for~~in the 2018 proposed rulemaking (as clarified in this SNPRM). The Agency continues to consider whether it is appropriate to rely on its authority in the above-reference environmental statutory provisions (potentially in conjunction with its housekeeping authority). The Agency will consider comments on this issue submitted in response to the ~~proposed rule.~~ 2018 proposal rulemaking and in response to this SNPRM.

II. Applicability to data and models

As identified by some public commenters, the 2018 proposed rulemaking did not put its discussion of increasing access to the data and models underlying pivotal regulatory science into the context of the broader approach that EPA uses to evaluate the entire body of scientific literature—that is, before it identifies candidates for "pivotal regulatory science." Under this regulation EPA would continue to use standard processes for identifying, evaluating, and reviewing available data, models, and studies. When the Agency has potentially identified multiple key studies or models of similar quality that could drive its subsequent decisions, the Agency will investigate the availability of the underlying data. If, for example, multiple high-quality studies exist but only two studies have data and models that are available for independent validation and reanalysis, EPA would only include those two studies as pivotal regulatory science and/or pivotal science in accordance with the 2018 proposed rulemaking. However, under the alternative approach in this supplemental proposal, EPA would consider using all available high-quality studies but give greater consideration to those two studies with

data available for independent validation.

As highlighted in some public comments, the terms "dose-response data and models," "dose-response models," "models" and "model assumptions" are not used consistently throughout the regulatory text of the 2018 proposed rulemaking. For example, some parts of the regulatory text appear to limit applicability of certain provisions to only dose-response models.[7] ~~.~~ In others, the requirements would apply more broadly ~~to all models.~~. EPA is now proposing a broader applicability. Transparency of EPA's science should not be limited to dose-response data and dose-response models, because other types of data and models will also drive the requirements and/or quantitative analysis of EPA final significant regulatory decisions and influential scientific information.

EPA is modifying the proposed regulatory text at 40 CFR 30.3, 30.5, 30.6 and 30.9 ~~to apply to all data and models~~ by deleting the term "dose-response," except in one instance. In proposed 40 CFR 30.6, EPA is not deleting "dose response" from the sentence specific to parametric dose-response models. ~~In addition,~~ EPA is ~~deleting~~also removing "including assumptions of a linear, no-threshold dose response" from 40 CFR 30.6~~.~~, because this could imply that the regulation is specific to those particular assumptions. This is not the intent of proposed 40 CFR 30.6.

~~In addition, where EPA refers only to "data" in the regulatory text at proposed § 30.5 of the 2018 proposed rule, EPA is now proposing to add the regulatory text "and models." This alternate approach to delete the regulatory term "dose-response" and add the regulatory term "and models" is identified as Alternate Option 1. As discussed~~

[7] See § 30.6

~~in Units I.C. and IV of this preamble, EPA is also considering two additional alternate approaches to proposed § 30.5.~~

Consistent with this broader approach to transparency, the proposed requirements of this rule would apply broadly to data and models underlying pivotal regulatory science and pivotal science which support significant regulatory decisions and influential scientific information, respectively, rather than simply to dose-response data and models. 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~~, and substantial risk reports.~~. The proposed definitions of "data" and "models" are discussed more fully in Section III.B of this preamble.

In addition, EPA is ~~also emphasizing~~clarifying that the ~~proposed requirement "EPA shall clearly explain the scientific basis for each model assumption used and present analyses showing the sensitivity~~use of the ~~modeled results to alternative~~ terms "model assumptions" and "models" in ~~§~~the proposed regulatory text at 40 CFR 30.6 ~~applies~~ apply to the assumptions that drive the model's analytic results, not to each model assumption used in the model~~, for example, not only chemical half-life but also body-weight.~~. EPA has modified the regulatory text at 40 CFR 30.6 to reflect this clarification.

EPA requests comment on the applicability of proposed 40 CFR 30.3, 30.5, 30.6 and 30.9 to ~~all~~ data and models.

III. Definitions

A. "Reanalyze" and "independent validation."

To improve the clarity of the proposed requirements, EPA is proposing definitions for certain terms.

In the 2018 proposed rulemaking, EPA used the terms "replicate" and "reproducible" and related terms. "Replicate" is used in the proposed regulatory text at 40 CFR 30.5. That section reads, in pertinent part, "[I]nformation is considered 'publicly available in a manner sufficient for independent validation' when it includes the information necessary for the public to understand, assess, and replicate findings…" "Replication" and "reproducibility" are both used in the 2018 proposed rulemaking preamble though neither is defined. Neither "reproducibility" nor its cognates are used in the regulatory text. "Replicate" was used but not defined in the regulatory text and its meaning was not discussed in the preamble.

Commenters contended that EPA was not clear about what it meant by the term "replicate" and that EPA appears to have conflated the term with "reproducible." Commenters interpreted the term "replicate" in several different ways. For example, some commenters contended that EPA used the term "replicate" but actually meant "reanalyze." EPA finds that these comments have merit and has determined that the intent of the term "replicate" should be clarified by establishing a regulatory definition.

EPA has considered the definitions in the National Academy of Sciences (NAS) "*Principles and Obstacles for Sharing Data from Environmental Health Research."* (Ref. 6, NAS Workshop Report), Pellizzari, et al. "*Reproducibility: A Primer on Semantics and Implications for Research"* (Ref. 7), and Goodman, et al. "*What does research*

reproducibility mean?" (Ref. 8). As demonstrated by these documents, there are varying definitions and understandings of these terms in the scientific community. Several commenters pointed to the use of the terms in the NAS Workshop Report (Ref. 6). The definitions in the NAS Workshop Report (Ref. 6) define "reanalysis," "replication," and "reproduce" as follows:

> A reanalysis is when you conduct a further analysis of data. A person doing a reanalysis of data may use the same programs and statistical methodologies that were originally used to analyze the data or may use alternative methodologies, but the point is to analyze exactly the same data to see if the same result emerges from the analysis.
>
> Replication means that you actually repeat a scientific experiment or a trial to obtain a consistent result. The second experiment uses exactly the same protocols and statistical programs but with different data from a different population. The goal is to see if the same results hold with data from a different population.
>
> When you reproduce, you are producing something that is very similar to that research, but it is in a different medium or context. In other words, a researcher who is reproducing an experiment addresses the same research question but from a different angle than the original researcher did.

EPA's use of "replicate" in the proposed regulatory text at 40 CFR 30.5 in the 2018 proposed rulemaking is generally consistent with the NAS Workshop Report (Ref. 6) definition of "reanalysis." However, as illustrated by Refs. 4-6, and in the public comments EPA received on the 2018 proposed rulemaking, these terms are not consistently used or defined in the scientific literature. EPA now proposes to use the term "reanalyze" instead of "replicate" in 40 CFR 30.5 and is clarifying the intent of the proposed regulation by proposing a definition of "reanalyze" at proposed 40 CFR 30.2 as "to analyze exactly the same data to see if the same result emerges from the analysis by using the same ~~programs and statistical methodologies that were originally used to analyze the data."~~ or different programs and statistical methodologies that were originally used to analyze the data." In addition to identifying potential analytical errors in the original work, reanalyzing the data would allow assessment of the robustness of

the original analysis and conclusions by, for instance, showing the variability that can occur when a previously omitted variable is added to the statistical model, different functional form assumptions are made (e.g., a linear marginal effect of treatment), or different assumptions are made when estimating standard errors and drawing statistical inferences (e.g., allowing for spatial correlation in error terms).

In the 2018 proposed rulemaking, EPA did not define "independent validation." The definition of "independent validation" depends on how the term "reanalyze" is defined. Independent validation for a scientific study that is being repeated by conducting a second scientific study would be different than independent validation where the data underlying a study is being reanalyzed to determine if the same results can be obtained. Thus, consistent with the proposed definition of "reanalyze" at proposed 40 CFR 30.2, EPA is proposing to define "independent validation" as the reanalysis of study data by subject matter experts who have not contributed to the development of the original study to demonstrate that the same analytic results are capable of being substantially reproduced. "Capable of being substantially reproduced" means 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.

EPA's interpretation of "capable of being substantially reproduced" as included in the proposed definition above builds on the description in the "*Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies"* (Ref. 9). These guidelines, which were issued by the Office of Management and Budget, are intended to help agencies ensure and maximize the

quality, utility, objectivity and integrity of the information that they disseminate ~~(*i.e.,* share with, or give access to, the public).~~.

EPA is requesting comment on the proposed definitions of "reanalyze" and "independent validation" at proposed 40 CFR 30.2.

B. Data and models.

Given the use of the term "data and models" in proposed 40 CFR 30.3, 30.5, 30.6 and 30.9 as described in Section II of this preamble, EPA is proposing to define "data" and "models" at proposed 40 CFR 30.2. EPA proposes to broaden the scope of the proposal by deleting the modifying term "dose-response," as indicated above, so as to extend the reference to data and models ~~to encompass *all* data~~underlying pivotal regulatory science and ~~models that are~~pivotal science used ~~in~~to support significant regulatory decisions and influential scientific information, respectively, not simply dose-response data and dose-response models. Examples of information that would be considered to be data and models for purposes of the proposed rulemaking 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~~, and substantial risk reports~~. This list is not exhaustive but is intended to provide examples of the range of information that would be considered to be within the scope of data and models.

1. *Data and research data.* Data has been defined to mean, in part, the recorded factual material commonly accepted in the scientific community as necessary to validate research findings (Ref. 10). As noted by public commenters and in the NAS Workshop

Report (Ref. 6), there are different stages of these data. "There are raw data, which come straight from the survey or the experiment. There are cleaned-up data, which consist of the raw data modified to remove obvious errors." (These are the data that are ready to be analyzed to extract relevant information.) "There are processed data, which are data that have been computed and analyzed to extract relevant information. There is the final clean data set that is provided with a publication. And there are the metadata that describe the data" (Ref. 6). These different data sets or stages of data may be used for different purposes and in different contexts.

The Agency received comment asking EPA to clarify what stage of data would need to be publicly available to allow for independent validation. Thus, EPA is incorporating the concept of stage of data with the ~~definition of research data~~basic concept of research data as "recorded factual material commonly accepted in the scientific community as necessary to validate research findings" from its definition at 2 CFR 200.315. For purposes of independent validation through reanalysis, the stage of data would be the cleaned-up or analyzable data set in which obvious errors have been removed. Obvious errors do not include data that could be characterized as outliers. These data are the ~~"~~cleaned-up or analyzable data set~~"~~ (Ref. 6). Therefore, EPA is proposing to define "data" as the set of recorded factual material commonly accepted in the scientific community as necessary to validate research ~~data~~findings in which obvious errors, such as keystroke or coding errors, have been removed and that is capable of being analyzed ~~to extract relevant information~~ by either the original researcher or an independent party. EPA requests comment on this proposed definition and whether the definition of "data" should apply to another stage of data described in

Ref. 6. The focus on this later stage of data is consistent with the Federal Government's approach to data access, and specifically to EPA's "*2016 Plan to Increase Access to Results of EPA-Funded Scientific Research"* (Ref. 11). Finally, EPA requests comment on whether there is another definition of "data" that should be considered.

EPA is deleting the 2018 proposed 40 CFR 30.2 definition of "research data," because this definition excludes "trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law" and "[p]ersonnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study." These types of data are intended to be subject to this rulemaking. To conform with this change, EPA is deleting the 2018 proposed 40 CFR 30.10 regulatory text because it would require EPA implementation of this rulemaking to be consistent with the definition of "research data."

2. Model. EPA is proposing to define "model" as it is defined in EPA's *Guidance on the Development, Evaluation, and Application of Environmental Model*s (Ref. 12). EPA's guidance document was produced to aid in strengthening the Agency's development, evaluation and use of models. In this guidance document, a model is described as "a simplification of reality that is constructed to gain insights into select attributes of a physical, biological, economic, or social system. A formal representation ~~of~~is characterized as the behavior of system processes, often in mathematical or statistical terms. The basis can also be physical or conceptual." This definition is based in part on the National Research Council's (NRC) 2007 report *Models in Environmental*

Regulatory Decision Making (Ref. 13). As noted by the NRC, models can be of many different forms. They can be computational, physical, empirical, conceptual or a combination of one or more types. EPA is requesting comment on the proposed definition of "model" at proposed 40 CFR 30.2.

~~EPA is requesting comment on the proposed definition of "model" at proposed § 30.2.~~

C. Publicly available.

In the 2018 proposed rulemaking, EPA used the term "publicly available" but did not define it at 40 CFR 30.2 or in the preamble to the 2018 proposed rulemaking. Given its use at 40 CFR 30.1, 30.5 and 30.9, EPA is proposing to define it. Publicly available information is often defined to mean information that is made available to the general public (e.g., see 40 CFR 2.201, 17 CFR 160.3, 16 CFR 313.3). EPA is proposing to define it similarly to how it is defined at 16 CFR 313.3. Although the overall purpose of the regulations at 16 CFR 313 is different than the purpose of this rulemaking, the meaning of information that is available to the general public should not vary. This would encompass information legally available from government sources, the media and the Internet. EPA is requesting comment on the proposed definition of "publicly available" at proposed 40 CFR 30.2.

IV. ~~Publicly Available~~Availability of Data and Models

In the 2018 proposed rulemaking, EPA proposed to require at 40 CFR 30.5 that "[w]hen promulgating significant regulatory decisions, the Agency shall ensure that dose-response data and models underlying pivotal regulatory science are publicly available in a manner sufficient for independent validation." ~~As discussed in Unit I.C.,~~

~~EPA proposes broadening this provision to include all data and models underlying pivotal regulatory science, rather than restricting the coverage of the provision only to just dose-response data and dose-response models.~~ EPA received a large number of comments stating that the approach in the 2018 proposed rulemaking would likely preclude the use of valid data and models from consideration as pivotal regulatory science, because the proposed requirement to make data and models publicly available in a manner sufficient for independent validation would prevent the use of data and models that include CBI data, proprietary data, PII data that cannot be ~~anonymized~~sufficiently de-identified to protect the data subjects, as well as many older studies. While ~~making~~having these data and models publicly available provides the greatest transparency, these commenters expressed concern that this approach could limit the use of certain data and models in EPA's significant regulatory decisions. ~~Thus, EPA is considering alternatives to this approach. EPA is proposing two alternate options that represent different approaches to the regulatory text at § 30.5 in addition to the proposed alternative described in Unit II of this preamble. Summaries of these alternate options are shown in Table 1.~~Based on a consideration of these comments, EPA is proposing a modified version of the 2018 proposed rulemaking regulatory text at 40 CFR 30.5. Proposed 40 CFR 30.5 would allow Agency consideration of studies where there is tiered access to data and models that have CBI, proprietary data, or PII that cannot be sufficiently de-identified to protect the data subjects. For all other studies, data and models should be publicly available if the studies are to be used as pivotal regulatory science or pivotal science.

~~TABLE 1-SUMMARY OF OPTIONS FOR PUBLICLY AVAILABLE DATA AND MODELS~~

Alternate Option	Scope of Application to Data and Models	Public Availability of Data and Models
1-All Data and Models Option[1]	All data and models regardless of when they were generated	All data and models must be publicly available
2-Weighing Option	All data and models regardless of when they were generated	Study is weighed based on whether data and models are publicly available
3-Tiered Data Access Option	All data and models regardless of when they were generated	Tiered approach to access to data and models that contain PII that cannot be anonymized, CBI and proprietary information. Access to CBI and proprietary information would be consistent with applicable statutes and regulations

[1] See Unit II.

A. Proposed § 30.5: Alternate Option 2 – Weighing Option.

This alternate option would weigh public availability with other assessment factors that characterize quality and relevance in determining whether data and models could be considered pivotal regulatory science. Thus, even if the data and models underlying pivotal regulatory science were not publicly available, *e.g.*, if a model's computer code were not available, EPA may consider a study as pivotal regulatory science if the other aspects of the data or model underlying the study were sufficiently robust. However, because all of the data and models would not be publicly available, EPA may assign lower weight to a study's evidence, findings and conclusions. It is crucial in making robust regulatory decisions that EPA has access to all aspects of data and models if they are to underly pivotal regulatory science. Thus, EPA would reserve the right to place less weight on the studies, to the point of entirely disregarding them, if the data and models underlying pivotal regulatory science are not made available in full to EPA.

The assessment factors that EPA may consider in conjunction with data availability are those that EPA currently considers when evaluating the quality and

~~relevance of scientific and technical information: soundness, applicability and utility, clarity and completeness, uncertainty and variability, and evaluation and review (Ref. 11). EPA requests comment on using these factors in conjunction with data availability in weighing a study. The proposed regulatory text for this alternate option includes definitions for each of these terms. EPA requests comment on whether these terms should be defined in § 30.2 rather than as part of § 30.5.~~

~~EPA is requesting comment on this alternate option for § 30.5.~~

~~EPA is also considering a variation of alternate option 2 because some of the data and models underlying pivotal regulatory science are older data and models that are not publicly available in a manner sufficient for independent validation. For example, the underlying data, models and computer code may not be readily publicly available because of the scientific norms for data and model availability that existed when they were developed. Given this consideration, EPA is requesting comment on whether this variation to alternate option 2 should apply only to data and models that are generated (*i.e.*, when the development of the data set or model has been completed or updated) after the effective date of this rulemaking. If this approach were finalized, EPA would weigh the lack of public availability only for data and models developed in the future. In this variation of alternate option 2, public availability of data and models sufficient for independent validation would not be a requirement for pivotal regulatory science generated before the effective date of the final rule. For this variation, EPA requests comment on whether the generation date of a model should be defined as the date on which the model or its underlying data were last updated.~~

~~Finally, EPA is requesting comment on whether the Agency should consider a~~

~~date earlier than the effective date of this rulemaking for purposes of this variation to alternate option 2, *i.e.*, a date some years prior to the publication of the 2018 proposed rule.~~

~~*B. Proposed § 30.5: Alternate Option 3 – Tiered Access Option.*~~

~~Alternate option 3 would still require that data and models be made available for independent validation but would not require that all data and models underlying pivotal regulatory science be *publicly* available. This option is consistent with the recent update to OMB's Information Quality Bulletin (Ref. 12~~ As discussed in OMB M-19-15 (Ref. 3), risk reduction techniques include creating multiple versions of a single dataset with varying levels of specificity and protection. The benefit of tiered access is that data users who wish to conduct activities with a statistical purpose without first obtaining special authorization have access to the versions of the data in the least restricted tiers, allowing them to conduct research while protecting confidentiality.

EPA is also proposing an alternative approach to today's proposed 40 CFR 30.5. Under alternative proposed 40 CFR 30.5, when promulgating significant regulatory decisions or finalizing influential scientific information, the Agency will, other things equal, give greater consideration to studies where the underlying data and models are available in a manner sufficient for independent validation either because the information is publicly available or available through tiered access when the data include CBI, proprietary data, or PII and appropriate techniques have been used to reduce the risk of re-identification. In developing the significant regulatory decision or influential scientific information, the EPA will identify those studies that are given greater consideration and provide a short description of why greater consideration was given.

However, the Agency may still consider studies where there is no access or limited access to underlying data and models.

EPA is also clarifying that the Agency does not intend to make all data and models underlying pivotal regulatory science and pivotal science publicly available. There may be instances where EPA does not own the data and models, lacks access to part or all of the data and models or does not have the authority to provide access to part or all of the data and models. Rather, EPA is describing how it will handle studies based on whether the underlying data and models are publicly available.

Both today's proposal and alternate proposal are consistent with EPA's statements in the April 2018 proposed rulemaking that "access to dose response data and models underlying pivotal regulatory science" should be done "in a manner consistent with statutory requirements for protection of privacy and confidentiality of research participants, protection of proprietary data and confidential business information, and other compelling interests" (Ref. 1). Both approaches are also based on EPA's recognition that there are statutory restrictions to public availability for some data and models that could make independent validation difficult. Further, both of these approaches are consistent with the OMB's M-19-15 (Ref. 3). OMB's implementation updates direct federal agencies to "explore methods that provide wider access to datasets while reducing the risk of disclosure of [PII]…[T]iered access offers promising ways to make data widely available while protecting privacy" (Implementation Update 3.5, Ref. 3). In addition, "Agencies should prioritize increased access to the data and analytic frameworks (*e.g.*, ~~models)"~~models) used to generate influential information" while being "consistent with statutory, regulatory, and policy requirements for

protections of privacy and confidentiality, proprietary data, and confidential business information" (Implementation Update 3.4, Ref. 3). This proposal is also consistent with OMB *Memorandum 13-13: Open Data Policy – Managing Information as an Asset* (Ref. 14).

~~It would provide for tiered access to data and models that have CBI, proprietary data, or PII that cannot be anonymized, and require that all other data and models be made publicly available if they are to be used as pivotal regulatory science. Under this option, studies could be used as pivotal regulatory science even if statutes or regulations restrict access to the data and models underlying those studies, provided that independent validation of these data and models could be accomplished in the restricted context by a person or persons authorized to have access to the restricted data and models. For data and models that do not include CBI, proprietary information, or PII that cannot be anonymized, this alternate option would still require that all such unrestricted data and models underlying pivotal regulatory science be publicly available. This option would balance decreased public access to the data and models underlying pivotal regulatory science with the continued use by EPA as pivotal regulatory science information that cannot be made publicly available.~~

Under a tiered approach to accessing data and models that include CBI, proprietary data, or PII that cannot be ~~anonymized,~~sufficiently de-identified to protect the data subjects, access is more restricted for more sensitive ~~the~~ data and models~~, the more restricted the access would be~~. Thus, the amount of information available for analysis is dictated by the tier. The greatest amount of information is made available at the most restricted access tier. Access to data involving PII would be consistent with

the requirements of the Common Rule, the Health Insurance Portability and Accountability Act (HIPPA), the 21st Century Cures Act, the Privacy Act, and other relevant laws and regulations, and EPA privacy policies. Reanalyzing findings of studies based on data and models that include PII (*e.g.*, residence) or CBI may not be possible given the degree of ~~redaction~~perturbation caused by de-identification that would be needed for the information to be made publicly available. Restricted access for researchers through secure data enclaves for PII or through non-disclosure agreements for CBI may result in access to sufficient information about the data and models to allow for independent validation. This ability to reanalyze findings may be much more limited for less restricted tiers. Thus, reanalysis of findings for some data and models may be limited to authorized researchers and not possible for the general public.

A model of tiered access ~~EPA is considering~~ for data involving PII[8] is the Research Data Center (RDC), National Center for Health Statistics (NCHS), Centers for Disease Control (CDC). The NCHS operates the RDC to allow researchers access to restricted-use data. The RDC provides access to the restricted-use data while protecting the confidentiality of survey respondents, study subjects, or institutions. For access to the restricted-use data, researchers must submit a research proposal outlining the need for restricted-use data. The submitted research proposal is intended to provide a framework for NCHS to identify potential disclosure risks and how the data will be used (Ref. 15). EPA is currently conducting a pilot study using the RDC's secure

~~[8] Access to data involving PII would be consistent with the requirements of the Common Rule (https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html), the Health Insurance Portability and Accountability Act (HIPPA), the 21st Century Cures Act, the Privacy Act, and other relevant laws and EPA privacy policies.~~

data enclave to host EPA datasets in a restricted use environment.

Development of standard data repositories is still ongoing. For example, the White House Office of Science and Technology Policy recently solicited public comments on a draft set of characteristics of data repositories used to locate, manage, share, and use data resulting from federally-funded research (85 FR 3085). The effort is intended to help federal agencies provide more consistent information on desirable characteristics of data repositories "for data subject to agency Public Access Plans and data management and sharing policies, whether those repositories are operated by government or nongovernmental entities." Information received during this public comment period will, among other things, help inform improved guidance and best practices related to preserving and providing access to data.

Access to CBI data would continue to be provided consistent with the environmental statutes EPA implements and the regulations at 40 CFR part 2, subpart B, which govern CBI. These regulations establish basic rules governing business confidentiality claims, how EPA handles business information that is or may be entitled to confidential treatment, and how EPA determines whether information is entitled to confidential treatment for reasons of business confidentiality. Various statutes under which EPA operates contain special provisions concerning the entitlement to confidential treatment of information gathered under such statutes. The regulations at 40 CFR part 2 subpart B prescribe rules for treating certain categories of business information obtained under the various statutory provisions.

~~Under this alternate option, for a subset of~~ In accordance with these statutes, both the proposed and alternative 40 CFR 30.5 provide that access to underlying data

and models ~~underlying pivotal regulatory science that~~ that include CBI, ~~PII or~~ proprietary information, ~~access to information on these data and models, in a manner sufficient to allow for independent validation of findings~~or PII, for the subset of studies that could be considered pivotal science, may be limited to authorized officials and researchers and not provided to the general public.

~~This alternate option maintains~~Proposed 40 CFR 30.5 would maintain the temporal approach to data and models taken in ~~§~~the regulatory text of 40 CFR 30.5 of the 2018 proposed ~~rule. This alternate proposal~~rulemaking, and thus would apply to ~~all data and models~~data and models evaluated at the time a significant regulatory action or influential scientific information is developed, regardless of when the data and models were generated. ~~-~~EPA is requesting comment on whether this should apply only to data and models that are generated (i.e., when the development of the data set or model has been completed or updated) after the effective date of this ~~alternate option~~rulemaking. 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.

Although the ability to independently validate pivotal regulatory science or pivotal science is a key component of this rulemaking, EPA would like to clarify that neither the proposed nor the alternative 40 CFR 30.5 would require that EPA, a member of the public or other entity must independently validate a study before it can be considered to be pivotal regulatory science or pivotal science. EPA would also like to clarify that independent validation is not required under proposed 40 CFR 30.7 which describes the

role of independent peer review.

EPA is requesting comment on the regulatory text being proposed today for ~~proposed § 30.5~~40 CFR 30.5. For alternate proposed 40 CFR 30.5, EPA is also requesting comment on how much consideration should be given to studies when there is limited or no access to the underlying data and models. In addition, EPA is requesting comment on how to ensure that, over time, more of the data and models underlying the science that informs significant regulatory decisions and influential scientific information are available to the public for independent validation in a manner that honors legal and ethical obligations to reduce the risks of unauthorized disclosure and re-identification. Finally, 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. Such comments will be used to develop implementation guidance.

V. Exemption by the Administrator

The 2018 proposed rulemaking includes a provision at 40 CFR 30.9 allowing 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 ~~all~~ data and models underlying pivotal regulatory science are publicly available in a ~~fashion~~manner that is consistent with law~~,~~ and protects privacy and confidentiality~~, and~~. EPA is ~~sensitive to national and homeland security (§ 30.9(a)), or in instances where OMB's Information Quality Bulletin for Peer Review provides for an~~ clarifying that the exemption ~~(§ 30.9(b)). Public commenters requested that EPA provide a clear, consistent, and systematic process for determining whether an exemption~~may be given

if compliance is ~~warranted under § 30.9(a). EPA has considered these comments and is modifying the regulatory text at § 30.9(a) to describe~~impracticable because technological barriers render sharing of the ~~approach EPA would take to recommend to the Administrator whether to grant an exemption.~~ data or models infeasible.

~~EPA proposes that the Administrator's decision whether to grant such an exemption be made based on a consideration of the following aspects of the study or studies that include data and models that cannot be made publicly available: soundness; applicability and utility; clarity and completeness; uncertainty and variability; evaluation and review (Ref. 11; Also see Unit IV.A. of this preamble); and as discussed below, the age of the study.~~

~~EPA requests comment on its proposed approach, including the criteria that the Administrator would consider in determining whether to grant an exemption under § 30.9(a).~~

EPA is also modifying the scope of the data and models that can be considered when determining whether to grant an exemption. ~~For older studies, as noted in Unit IV.A, the~~The underlying data, models and computer code for some studies, particularly older studies, may not be readily publicly available because of the ~~scientific norms for~~ technological barriers to data and model ~~availability~~sharing (e.g., differences in data storage devices or data retention practices) that existed when they were developed. Thus, in 40 CFR 30.9(a~~) as part of~~), EPA is proposing to use the age of data and models as a factor in the determination that compliance with the rule is impracticable ~~because it is not feasible to ensure that data and models underlying pivotal regulatory science are publicly available in a manner sufficient~~. This modification of scope is

intended to acknowledge the evolution of best practices for ~~independent validation, EPA is proposing the age of data and models as a factor for consideration~~information sharing given innovations in ~~making that determination.~~information generation, access, management and use (See Ref. 3). EPA is proposing that ~~when the development of data or models was completed or updated before the effective date of this rule, the~~a study or studies would be eligible for consideration under 40 CFR 30.9(a), regardless of whether they contain CBI, proprietary information~~.~~, or PII, if the underlying data or models were collected, completed or updated before the effective date of this rule. EPA requests comment on this consideration of the age of data and models in determining the feasibility of making underlying data and models publicly available. EPA also requests comment on whether ~~it should consider other~~ there are aspects ~~of the scientific norms for data and model availability rather~~ other than the year the data or model was collected, completed or updated that EPA should consider in determining whether to grant an exemption in order to evaluate the technological barriers to sharing.

The 2018 proposed rulemaking also included a provision at 40 CFR 30.9 allowing 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 conduct independent peer review on all pivotal regulatory science. EPA is deleting that provision of the proposed exemption because EPA does not believe that peer review of pivotal regulatory science or pivotal science would be infeasible. Thus, EPA no longer believes the provision is necessary.

VI. References

The following is a listing of the documents that are specifically referenced in this

notice. The docket includes these documents and other information considered by EPA, including documents referenced within the documents that are included in the docket, even if the referenced document is not physically located in the docket. For assistance in locating these other documents, please consult the person listed under **FOR FURTHER INFORMATION CONTACT**.

1. EPA. Strengthening Transparency in Regulatory Science; Proposed Rule. **Federal Register** (83 FR18768, April 30, 2018) (FRL-9977-40).

2. EPA. Strengthening Transparency in Regulatory Science; Extension of Comment Period and Notice of Public Hearing **Federal Register** (83 Fed. Reg. 24255, May 25, 2018).

3. OMB (Office of Management and Budget). (2019). Improving Implementation of the Information Quality Act. Memorandum for the Heads of Executive Departments and Agencies. OMB Issuance M-19-15. Washington, DC: Executive Office of the President. https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf

4. OMB (Office of Management and Budget). (2004). Memorandum for the Heads of Executive Departments and Agencies on Final Information Quality Bulletin for Peer-Review. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/memoranda/fy2005/m05-03.pdf.

5. OMB (Office of Management and Budget). (2019). Executive Order 13891 of October 9, 2019. Promoting the Rule of Law Through Improved Agency Guidance Documents. 84 FR 199. https://www.govinfo.gov/content/pkg/FR-2019-10-15/pdf/2019-22623.pdf.

6. NAS (National Academies of Sciences, Engineering, and Medicine). (2016). Principles and obstacles for sharing data from environmental health research: Workshop summary. Washington, DC: The National Academies Press. https://doi.org/10.17226/21703.

7. Pellizzari, YE; Lohr, K, Blatecky, A.; Creel, D. (2017). Reproducibility: A Primer on Semantics and Implications for Research. Research Triangle Park, NC: RTI Press.

8. Goodman, SN; Fanelli, D; Ioannidis, JPA. (2016). What does research reproducibility mean? Sci Translational Medicine 8: 341ps12. https://doi.org/10.1126/scitranslmed.aaf5027

9. OMB (Office of Management and Budget). (2002). Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Final guidelines. 67 FR 8452-8460. https://www.govinfo.gov/content/pkg/FR-2002-02-22/pdf/R2-59.pdf

10. OMB (Office of Management and Budget). (2013). Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards; Final Rule. 78 FR 78589-78691. https://www.govinfo.gov/content/pkg/FR-2013-12-26/pdf/2013-30465.pdf

11. U.S. EPA (U.S. Environmental Protection Agency). (2016). Plan to Increase Access to Results of EPA-Funded Scientific Research. (EPA/601-R-16-005). Washington, DC: U.S. Environmental Protection Agency. https://www.epa.gov/sites/production/files/2016-12/documents/epascientificresearchtransperancyplan.pdf

12. U.S. EPA (U.S. Environmental Protection Agency). (2009). Guidance on the Development, Evaluations, and Application of Environmental Models. (EPA/100/K-

09/003). Washington, DC: US. Environmental Protection Agency. https://www.epa.gov/sites/production/files/2015-04/documents/cred_guidance_0309.pdf

~~11. U.S EPA (U.S. Environmental Protection Agency). (2014). Framework for Human Health Risk Assessment to Inform Decision Making. (EPA/100/R-14/001). Washington, DC: U.S. Environmental Protection Agency, Office of the Science Advisor, Risk Assessment Forum. https://www.epa.gov/sites/production/files/2014-12/documents/hhra-framework-final-2014.pdf~~

~~12.~~ 13. NRC (National Research Council). (2007). Models in Environmental Regulatory Decision Making. Washington, DC: The National Academies Press. ~~13~~https://doi.org/10.17226/11972

14. OMB (Office of Management and Budget). (2013). Memorandum for the Heads of Executive Departments and Agencies on Open Data Policy—Managing Information as an Asset (https://projectopen-data.cio.gov/policy-memo/).

15. CDC (Centers for Disease Control). Research Data Center. https://www.cdc.gov/rdc/index.htm

VII. Statutory and Executive Order Reviews

Additional information about these statutes and Executive Orders can be found at https://www.epa.gov/laws-regulations/laws-and-executive-orders.

A. Executive Order 12866: Regulatory Planning and Review and Executive Order 13563: Improving Regulation and Regulatory Review

This action is a significant regulatory action that was submitted to the Office of Management and Budget (OMB) for review. Any changes made in response to OMB recommendations have been documented in the docket.

B. Executive Order 13771: Reducing Regulations and Controlling Regulatory Costs

This action is not expected to be an Executive Order 13771 regulatory action because it relates to ‘‘agency organization, management or personnel.’’

C. Paperwork Reduction Act (PRA)

This action does not contain any information collection activities and therefore does not impose an information collection burden under the PRA.

D. Regulatory Flexibility Act (RFA)

I certify that this action will not have a significant economic impact on a substantial number of small entities under the RFA. This action will not impose any requirements on small entities. This action does not regulate any entity outside the federal government.

E. Unfunded Mandates Reform Act (UMRA)

This action does not contain any unfunded mandate as described in UMRA, 2 U.S.C. 1531–1538, and does not significantly or uniquely affect small governments. The action imposes no enforceable duty on any state, local or tribal governments or the private sector.

F. Executive Order 13132: Federalism

This action does not have federalism implications. It will not have substantial direct effects on the states, on the relationship between the national government and the states, or on the distribution of power and responsibilities among the various levels of government.

G. Executive Order 13175: Consultation and Coordination With Indian Tribal Governments

This action does not have tribal implications as specified in Executive Order 13175. Thus, Executive Order 13175 does not apply to this action.

H. Executive Order 13045: Protection of Children From Environmental Health Risks and Safety Risks

The EPA interprets Executive Order 13045 as applying only to those regulatory actions that concern environmental health or safety risks that the EPA has reason to believe may disproportionately affect children, per the definition of ‘‘covered regulatory action’’ in section 2–202 of the Executive Order. This action is not subject to Executive Order 13045 because it does not concern an environmental health risk or safety risk.

I. Executive Order 13211: Actions Concerning Regulations That Significantly Affect Energy Supply, Distribution or Use

This action is not a ‘‘significant energy action’’ because it is not likely to have a significant adverse effect on the supply, distribution or use of energy.

J. National Technology Transfer and Advancement Act (NTTAA)

This rulemaking does not involve technical standards.

K. Executive Order 12898: Federal Actions To Address Environmental Justice in Minority Populations and Low-Income Populations

The EPA believes that this action is not subject to Executive Order 12898 (59 FR 7629, February 16, 1994) because it does not establish an environmental health or safety standard.

List of Subjects in 40 CFR Part 30

Environmental protection, Administrative practice and procedure, Reporting and recordkeeping requirements

Strengthening Transparency in Regulatory Science

Page XX of XX

Dated: _____.

_____

Andrew R. Wheeler, Administrator.

Therefore, 40 CFR part 30, as proposed to be added at 83 FR 18768 (April 30, 2018), is proposed to be amended as follows:

PART 30—TRANSPARENCY IN REGULATORY DECISIONMAKING

1. The authority citation for part 30 is revised to read as follows:

Authority: ~~Departmental Regulations.~~ 5 U.S.C. 301; 5 U.S.C. App.; Pub. L. 98-80, 84 Stat. 2086

2. Revise § 30.2 to add the following:

§ 30.2. What definitions apply to this subpart?

"*Capable of being substantially reproduced*" means 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.

Data means the set of recorded factual material commonly accepted in the scientific community as necessary to validate research ~~data~~findings in which obvious errors, such as keystroke or coding errors, have been removed and that is capable of being analyzed ~~to extract relevant information~~ by either the original researcher or an independent party.

Influential scientific information means scientific information the agency reasonably can determine will have or does have a clear and substantial impact on important public policies or private sector decisions.

Independent validation means the reanalysis of study data by subject matter experts who have not contributed to the development of the study to demonstrate that the same analytic results reported in the study are capable of being substantially reproduced.

Model means a simplification of reality that is constructed to gain insights into select

attributes of a physical, biological, economic, or social system. A formal representation of the behavior of system processes, often in mathematical or statistical terms. The basis can also be physical or conceptual.

Pivotal science means the specific scientific studies or analyses that underly influential scientific information.

Publicly available means lawfully available to the general public from federal, state, or local government records; the Internet; widely distributed media; or disclosures to the general public that are required to be made by federal, state, or local law.

Reanalyze means to analyze exactly the same data to see if the same result emerges from the analysis by using the same ~~programs~~or different statistical software, models, and statistical methodologies that were originally used to analyze the data, as well as to assess potential analytical errors and variability in the underlying assumptions of the original analysis.

3. Revise § 30.3 to read as follows:

§ 30.3 How do the provisions of this subpart apply?

The provisions of this subpart apply to data and models**,** underlying pivotal science supporting influential scientific information and/or underlying pivotal regulatory science ~~that are~~ used to justify significant regulatory decisions regardless of the source of funding or identity of the party conducting the ~~regulatory~~ science. The provisions of this section do not apply to physical objects (like laboratory samples), drafts, and preliminary analyses. In the event the procedures outlined in this rule conflict with statutes that EPA administers, or their implementing regulations, the statutes and regulations will control. Except where explicitly stated otherwise, the provisions of this

subpart do not apply to any other type of agency action, including individual party adjudications, enforcement activities, or permit proceedings.

§ 30.5 [Reserved]

4. Revise § 30.5 to read as follows:

§ 30.5 What requirements apply to EPA's use of data and models underlying pivotal regulatory science~~?~~ and pivotal science?

When promulgating significant regulatory decisions~~,~~ or finalizing influential scientific information, the Agency ~~shall ensure that data and models underlying~~ will only use pivotal regulatory science and/or pivotal science that includes 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, and studies that do not include restricted data and models if the data and models are publicly available in a manner sufficient for independent validation.

Where the Agency is making data or models publicly available, it shall do so in a ~~fashion~~manner that is consistent with law, protects privacy, confidentiality, confidential business information, and is sensitive to national and homeland security. Information is considered ''~~publicly~~ available in a manner sufficient for independent validation'' when it includes the information necessary ~~for the public~~ to understand, assess, and reanalyze findings. This may include, for example:

(a) Data (where necessary, data would be made available subject to access and use restrictions);

(b) Associated protocols necessary to understand, assess, and extend conclusions;

(c) Computer codes and models involved in the creation and analysis of such information;

(d) Recorded factual materials; and

(e) Detailed descriptions of how to access and use such information.

The provisions of this section apply to data and models~~,~~ underlying pivotal regulatory science or pivotal science regardless of who funded or conducted the underlying data, models, or other regulatory science or pivotal science. The agency shall make ~~all~~ reasonable efforts to explore methodologies, technologies, and institutional arrangements for making such data and models available before it concludes that doing so in a manner consistent with law and protection of privacy, confidentiality, national and homeland security is not possible. Where data ~~is~~and models are controlled by third parties, EPA ~~shall~~may work with those parties to endeavor to make the data and models available in a manner that complies with this section.

5. Revise § 30.5 to read as follows:

§ 30.5 What requirements apply to EPA's use of data and models underlying pivotal regulatory science~~?~~ and pivotal science? [ALTERNATE]

~~When promulgating significant regulatory decisions, the Agency shall give higher priority to studies based on data and models that are publicly available in a manner sufficient for independent validation. When the data and models underlying pivotal regulatory science are not made available to EPA or are not publicly available in a manner sufficient for independent validation because they include confidential business information, proprietary data and models or personally identifiable information, EPA may~~

~~use the data and models but may assign lower weight to the studies' evidence, findings and conclusions. Factors that EPA may consider in determining the weight to assign to these studies may include soundness, applicability and utility, clarity and completeness, uncertainty and variability, and evaluation and review. EPA would reserve the right to place less weight on the studies, to the point of entirely disregarding them, if the data and models underlying those studies are not made available in full to EPA. Where the Agency is making data or models publicly available, it shall do so in a fashion that is consistent with law, protects privacy, confidentiality, confidential business information, and is sensitive to national and homeland security. Information is considered "publicly available in a manner sufficient for independent validation" when it includes the information necessary for the public to understand, assess, and reanalyze findings. This may include, for example:~~

~~(a) Data (where necessary, data would be made available subject to access and use restrictions);~~

~~(b) Associated protocols necessary to understand, assess, and extend conclusions;~~

~~(c) Computer codes and models involved in the creation and analysis of such information;~~

~~(d) Recorded factual materials; and~~

~~(e) Detailed descriptions of how to access and use such information.~~

~~Soundness is the extent to which the scientific and technical procedures, measures, methods or models employed to generate the information are reasonable for, and consistent with, the intended application.~~

~~Applicability and Utility is the extent to which the information is relevant for the intended use.~~

~~Clarity and Completeness is the degree of clarity and completeness with which the data, assumptions, methods, quality assurance, and analyses employed to generate the information are documented.~~

~~Uncertainty and Variability is the extent to which the variability and uncertainty (quantitative and qualitative) in the information or in the procedures, measures, methods or models are evaluated and characterized.~~

~~Evaluation and Review is the extent of independent verification, validation and peer review of the information or of the procedures, measures, methods or models.~~

~~The provisions of this section apply to data and models underlying pivotal regulatory science regardless of who funded or conducted the underlying data, models, or other regulatory science. The agency shall make all reasonable efforts to explore methodologies, technologies, and institutional arrangements for making such data and models available before it concludes that doing so in a manner consistent with law and protection of privacy, confidentiality, national and homeland security is not possible. Where data and models are controlled by third parties, EPA shall work with those parties to endeavor to make the data and models available in a manner that complies with this section.~~

~~[PROPOSED REGULATORY TEXT FOR § 30.5 ALTERNATE OPTION 3]~~

~~§ 30.5 What requirements apply to EPA's use of data and models underlying pivotal regulatory science?~~

~~When promulgating significant regulatory decisions, the Agency shall ensure that data and models underlying pivotal regulatory science, except data and models that include confidential business information, proprietary information or personally identifiable information, are publicly available in a manner sufficient for independent validation. Studies~~ When promulgating significant regulatory decisions or finalizing influential scientific information, the Agency will, other things equal, give greater consideration to studies where the underlying data and models are publicly available in a manner sufficient for independent validation. The Agency will also give greater consideration to studies based on data and models that include confidential business information, proprietary information or personally identifiable information ~~could be considered to be pivotal regulatory science~~ if these data and models were available through restricted access, such as through a secure data enclave, in a manner sufficient for independent validation. Where there is no access to data and models, or access is limited, the Agency may still consider these studies, depending on the other attributes of the studies. Furthermore, the Agency will identify those studies that are given greater consideration and provide a short description of why greater consideration was given. Where the Agency is making data or models publicly available, it shall do so in a ~~fashion~~manner that is consistent with law, protects privacy, confidentiality, confidential business information, and is sensitive to national and homeland security. Information is considered ‘‘available in a manner sufficient for independent validation’’ when it includes the information necessary to understand, assess, and reanalyze findings. This may include, for example:

(a) Data (where necessary, data would be made available subject to access and use

restrictions);

(b) Associated protocols necessary to understand, assess, and extend conclusions;

(c) Computer codes and models involved in the creation and analysis of such information;

(d) Recorded factual materials; and

(e) Detailed descriptions of how to access and use such information.

The provisions of this section apply to data and models underlying pivotal regulatory science or pivotal science regardless of who funded or conducted the underlying data, models, or other regulatory science~~.~~ or pivotal science. The agency shall make ~~all~~ reasonable efforts to explore methodologies, technologies, and institutional arrangements for making such data and models available before it concludes that doing so in a manner consistent with law and protection of privacy, confidentiality, national and homeland security is not possible. Where data and models are controlled by third parties, EPA ~~shall~~may work with those parties to endeavor to make the data and models available in a manner that complies with this section.

6. Revise § 30.6 to read as follows:

§ 30.6 What additional requirements pertain to the use of data and models underlying pivotal science or pivotal regulatory science?

EPA shall describe and document any assumptions and methods used and shall describe variability and uncertainty. EPA shall evaluate the appropriateness of using default assumptions on a case-by-case basis. EPA shall clearly explain the scientific basis for ~~each model assumption used and~~critical assumptions used in the analysis that drove the analytical results and subsequent decisions and shall present analyses

showing the sensitivity of the modeled results to alternative assumptions. When available, EPA shall give explicit consideration to high quality studies, including but not limited to those that explore: A broad class of parametric dose-response or concentration-response models; a robust set of potential confounding variables; nonparametric models that incorporate fewer assumptions; various threshold models across the dose or exposure range; and models that investigate factors that might account for spatial heterogeneity.

7. Revise § 30.7 to read as follows:

§ 30.7 What role does independent peer review have in this section?

EPA shall conduct independent peer review on all pivotal regulatory science used to justify significant regulatory decisions and on all pivotal science underlying influential scientific information, consistent with the requirements of the OMB Final Information Quality Bulletin for Peer Review (70 FR 2664) and the exemptions described therein. Because transparency in regulatory science includes addressing issues associated with assumptions used in models, EPA shall ask peer reviewers to articulate the strengths and weaknesses of EPA's justification for the assumptions applied and the implications of those assumptions for the results.

8. Revise § 30.9 to read as follows:

§ 30.9 May the EPA Administrator grant exemptions to this subpart?

The Administrator may grant an exemption to this subpart on a case-by case basis if he or she determines that compliance is impracticable because~~:~~

~~(a) It is not feasible to ensure that all data and~~ technological barriers render sharing of the data or models ~~underlying pivotal regulatory science are publicly available in a~~

~~manner sufficient for independent validation because~~infeasible, the development of the data or model was completed or updated before [*EFFECTIVE DATE OF FINAL RULE*] or making the data and models publicly available would conflict with laws governing privacy, confidentiality, confidential business information, or national and homeland security~~; or~~

~~(b) It is not feasible to conduct independent peer review on all pivotal regulatory science used to justify regulatory decisions for reasons outlined in OMB Final Information Quality Bulletin for Peer Review (70 FR 2664). Section IX.~~

~~In determining whether the study or studies should be exempt under subsection (a), the Administrator shall consider the following aspects of the study or studies that include the data and models that cannot be made publicly available.~~

~~Soundness. The extent to which the scientific and technical procedures, measures, methods or models employed to generate the information are reasonable and consistent with the intended application.~~

~~Applicability and utility. The extent to which the information is relevant for the intended use.~~

~~Clarity and completeness. The degree of clarity and completeness with which the data, assumptions, methods, quality assurance and analyses employed to generate the information are documented.~~

~~Uncertainty and variability. The extent to which the variability and uncertainty (quantitative and qualitative) in the information or in the procedures, measures, methods or models are evaluated and characterized.~~

~~Evaluation and review. The extent of independent verification, validation and~~

~~peer review of the information, procedures, measures, methods or models.~~

~~Whether the development of the data or model was completed or updated before [insert effective date].~~.