FDA's Response to Peer Review Comments on FDA's Post-Market Assessment Prioritization Tool

May 2026

TABLE OF CONTENTS

I. INTRODUCTION

Versar Global Solutions (Versar), an independent U.S. Food and Drug Administration (FDA) contractor, coordinated an external peer review of FDA's Post-market Assessment Prioritization Tool.

Versar contacted fifteen experts to fulfill FDA's requests for: one to two experts in Multi-Criteria Decision Analysis (MCDA), risk assessment, decision analysis, and/or decision-making in government; two to four experts in toxicology and/or exposure assessment of food chemicals including authorized substances or contaminants; and one to two experts in epidemiology and/or public health. As a result, Versar found five experts who were interested and available to participate and spanned the necessary expertise. For each interested and available peer reviewer, Versar conducted conflict of interest (COI) screening to ensure that the experts had no real or perceived COI.

Peer Reviewers:

Barbara Engeli, MSc

Ms. Engeli has twenty years of experience in chemical risk assessment/regulatory human toxicology and food safety at the Federal Food Safety and Veterinary Office and Federal Office of Public Health in Switzerland. Ms. Engeli has conducted chemical risk assessments on food and commodities, in particular contaminants, novel food, botanicals, food supplements, veterinary drug residues, cosmetics, and toys. Ms. Engeli holds a Master of Science in Biology from the University of Zurich.

Mary A. Fox, PhD, MPH

Dr. Fox is an Associate Practice Professor at the Johns Hopkins University School of Public Health in the Departments of Health Policy and Management and Environmental Health and Engineering. She develops and applies methods for cumulative impact and risk assessment to reduce environmental health disparities, design and lead community and occupational health investigations, and policy and program evaluations. Dr. Fox holds a PhD from Johns Hopkins University Bloomberg School of Public Health, Environmental and Occupational Health Policy.

George M. Gray, PhD

Dr. Gray is a professor in the Department of Environmental and Occupational Health at George Washington University (GWU) Milken Institute School of Public Health. He is Director for the Center for Risk Science and Public Health at GWU and Associate Dean for MPH Programs. Dr.

Gray holds a PhD in Toxicology from the University of Rochester School of Medicine and Dentistry.

Norbert E. Kaminski, PhD

Dr. Kaminski is Endowed Chair for the Food and Consumer Product Ingredient Safety, Director for the Center for Research on Ingredient Safety, Director the Institute for Integrative Toxicology, and Professor of Pharmacology and Toxicology in the Cell and Molecular Biology Program at Michigan State University. His lab conducts research focused on elucidation of the molecular mechanisms for impairment of signal transduction cascades and gene expression during lymphocyte activation by drugs and chemicals. Dr. Kaminski holds a PhD in Toxicology and Physiology from North Carolina State University.

Salomon Sand, PhD

Dr. Sand is a Risk Assessor for the Department of Risk-Benefit Assessment at the Swedish Food Agency. He coordinates development of risk and benefit assessment methodology at the department across the Toxicology, Biohazards, and Nutrition units. Dr. Sand holds a Medical PhD in Toxicology and quantitative health risk assessment from the Karolinska Institute, Sweden.

II. CHARGE TO REVIEWERS

The focus of this review is the draft Post-market Assessment Prioritization Tool, which is a critical part of FDA's overall systematic post-market assessment process. Comments from the technical reviewers will be used to inform further development and refinement of the Post-market Assessment Prioritization Tool. The intention is for the reviewers to provide a technical review of the *Post-market Assessment Prioritization Tool* by answering the charge questions. The FDA posted the tool and questions for public comment on June 18, 2025; the docket closed August 18, 2025. We request the experts review these comments before answering the charge questions. A description of the tool and the charge questions were posted here (https://www.regulations.gov/document/FDA-2025-N-1733-0001).

It should be noted that since the tool was sent out for public comment in June 2025, FDA has continued to develop and refine the systematic post-market assessment process. As a result, several changes have been made to the process described in the section of the prioritization tool titled "Overview". An updated description of the post-market assessment process and how the prioritization tool would be applied within that process is described below.

The systematic post-market assessment of food chemicals consists of the following three phases and will be published in more detail later this year:

1. Review of Information (signal detection and triage; prioritization; and finalization of an annual workplan);
2. Scientific Assessment (safety, risk, and/or hazard); and
3. Risk Management

Phase 1, Review of Information, includes three steps: signal detection and triage; prioritization; and finalization of an annual workplan. The prioritization tool is FDA's proposed method of prioritizing chemicals to be added to the annual workplan. Food chemicals added to the yearly workplan move into Phase 2, of the systematic post-market assessment process, Scientific Assessment, and are subject to comprehensive and multi-step evaluations. These assessments will consider all scientific evidence available to FDA to establish robust scientific conclusions on food chemical safety.

Although chemical prioritization is only one part of the larger post-market systematic assessment process, it provides a framework that assists FDA in allocating limited resources to the food chemicals with the most urgent need for scientific assessment. FDA recognizes that the chemical prioritization process must be science-based, data-driven, systematic, and reproducible. However, the chemical prioritization process must also be rapid such that the annual post-market workplan reflects the most current list of food chemicals for scientific assessment.

Charge Questions

FDA seeks to develop a science-based, data-driven, systematic, and reproducible process for the prioritization of chemicals in food that are candidates for post-market assessments, including food additives; color additives; generally recognized as safe (GRAS) substances (including GRAS substances that have not been notified to FDA); food contact substances; and chemicals that are present as unintentional contaminants. The draft Post-market Assessment Prioritization Tool focuses on potential risk to public health (risk ranking) and also includes other decisional considerations, using a Multi-Criteria Decision Analysis (MCDA) method.

The focus of this review is the draft Post-market Assessment Prioritization Tool, which is a critical part of FDA's overall systematic post-market assessment process. Considering the information in Sections 1-3 and the appendix of this document, please provide feedback on the following questions:

The purpose of the Post-market Assessment Prioritization Tool is to assist in making decisions about which chemicals, including both intentionally added substances and unintentional contaminants in food, are a priority to review. Is the modeling approach we proposed appropriate for this purpose? If not, please explain your reasoning and provide alternatives for FDA to consider. Please be specific and provide references, as appropriate.

1. The purpose of the Post-market Assessment Prioritization Tool is to assist in making decisions about which chemicals, including both intentionally added substances and unintentional contaminants in food, are a priority to review. Is the modeling approach we proposed appropriate for this purpose? If not, please explain your reasoning and provide alternatives for FDA to consider. Please be specific and provide references, as appropriate.

2. The draft Post-market Assessment Prioritization Tool currently includes four Public Health criteria and three Other Decisional criteria.

a. Are the four Public Health criteria appropriate for the purpose of the tool? If not, please explain what changes might be considered and why.
b. Are the three Other Decisional criteria appropriate for the purpose of the tool? If not, please explain what changes might be considered and why.
c. Are there additional criteria that should be considered? If so, please describe additional criteria that might be considered and why.

3. The draft scoring definitions for all criteria were developed to consider the expected variability in the types and extent of data available for the wide variety of food chemicals that may be considered for review.

a. Given this context, are the scoring definitions for the Public Health criteria appropriate for the purpose of the tool?
    i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why?
    ii. The toxicity criterion described in Section 3.1.1 considers data for seven different toxicity data types and the score assigned reflects

the highest toxicity data type score from the toxicity rubric, which is described in Appendix A Table A1. Is this the most appropriate strategy for assigning a toxicity criterion score? If not, please explain your reasoning and provide alternatives for FDA to consider. Please be specific and provide references, as appropriate.

b. Are the scoring definitions for the Other Decisional criteria appropriate for the purpose of the tool?
    i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why.
    ii. FDA is exploring quantitative and qualitative methods to help inform the scoring of the 'building public confidence' criterion (Section 3.2.3) such as conducting public sentiment analysis (e.g., utilizing natural language processing). How might such tools or the information they provide be incorporated into this criterion? What additional strategies and metrics could FDA consider?

4. The prioritization methodology includes weighting factors.

a. FDA is considering equal weighting among the Public Health criteria and (separately), among the Other Decisional criteria for the Post-market Assessment Prioritization Tool.
    i. Should different weights be applied to the Public Health criteria when determining the Total Public Health Criteria Score? If so, please specify the weighting scheme that might be considered and why.
    ii. Should different weights be applied to the Other Decisional Criteria when determining the Total Other Decisional Criteria Score? If so, please specify the weighting scheme that might be considered and why.
b. FDA is considering equal weighting among the Total Public Health Criteria Score and the Total Other Decisional Criteria Score to determine the overall Post-market Assessment Prioritization Score.
    i. Should different weights be applied when determining the overall Post-market Assessment Prioritization Score? If so, please specify the weighing scheme that might be considered and explain why it would be more appropriate than equal weighting.

5. The draft toxicity rubric uses traditional toxicity data *(in vivo*, as well as limited *in vitro* such as for genotoxicity), human health outcomes (e.g., adverse event reports), and epidemiological data for determination of the toxicity criterion score within the Public Health criteria. Considering that the prioritization process is not a comprehensive review, please address the following questions.

a. How might FDA incorporate information from new approach methodologies (NAMs) into the toxicity rubric?
    i. Are there specific NAMs (e.g., systems biology, engineered tissues, artificial intelligence, *in vitro*, microphysiological systems, or other alternative data or modeling tools) that would be most appropriate for use in the toxicity rubric? If so, please explain which NAM(s) would be most appropriate and why.

ii. Given that a single NAM is not expected to be a one-to-one replacement for a traditional *in vivo* toxicity test, how can the strengths and limitations of each NAM be appropriately considered if it is incorporated into the toxicity rubric?

b. Threshold of Toxicological Concern (TTC) approaches can be used to assess the toxicity of chemicals that lack sufficient safety data and have low dietary exposures. Although the Cramer classification scheme has historically been used in TTC approaches, FDA has recently developed the Expanded Decision Tree (EDT) that assigns chemicals to one of six EDT classes. How might such tools or the information they provide be incorporated into the toxicity rubric?

6. Do you have any additional comments? Please share them in your review.

III. SUMMARY OF PEER REVIEWER COMMENTS AND FDA'S RESPONSE

Question 1: The purpose of the Post-market Assessment Prioritization Tool is to assist in making decisions about which chemicals, including both intentionally added substances and unintentional contaminants in food, are a priority to review. Is the modeling approach we proposed appropriate for this purpose? If not, please explain your reasoning and provide alternatives for FDA to consider. Please be specific and provide references, as appropriate.

There was a consensus among reviewers that the modeling approach is appropriate. Reviewers also provided feedback in this section regarding other aspects of the tool and FDA's post-market systematic process.

Two reviewers commented that it would be helpful to know more about how the prioritization tool fits into the overall systematic process, especially the steps preceding prioritization (i.e., signal detection and triage) and how those steps impact whether prioritization happens. One reviewer suggested providing a flowchart of the overall systematic process.

Two reviewers commented about the importance of conducting the prioritization and the overall systematic process with transparency.

One reviewer commented that it would be helpful to know more about the prior round of pilot testing and lessons learned from that.

Two reviewers suggested there be two separate lists of prioritized chemicals with one for intentionally added substances and one for unintentional contaminants.

Some reviewers provided other comments/suggestions regarding criteria, scoring definitions, and weighting which they also provided (i.e., repeated) in questions specific to these respective topics.

FDA Response:

We thank the reviewers for their thorough review and comments and/or suggestions regarding the appropriateness of the modeling approach, as well as other more general comments/suggestions provided here.

The revised prioritization tool document is published on the FDA's "Understanding the FDA's Systematic Process for Ensuring the Post-Market Safety of Chemicals in Food" webpage[1] and is also accessible via a link on the *Completed Reviews* section of the *Peer Review of Scientific Information and Assessments* webpage.[2] Published along with the revised prioritization tool document is *FDA's Update on an Enhanced Systematic Process for FDA's Post-Market Assessment of Chemicals in Food* (i.e., companion paper) which provides detailed information about the overall systematic process,

[1] https://www.fda.gov/food/food-chemical-safety/understanding-fdas-systematic-process-ensuring-post-market-safety-chemicals-food
[2] https://www.fda.gov/science-research/peer-review-scientific-information-and-assessments/completed-peer-reviews

including signal detection and triage, prioritization, scientific assessment, and risk management. A process diagram is included in the companion paper that provides a graphic representation of the systematic post-market assessment process from beginning to end.

Following prioritization, food chemicals selected for review will undergo a full scientific assessment. These scientific assessments will be made publicly available because they are critical to stakeholders -- they inform safety determinations and risk management plans, which may lead to regulatory decisions. In contrast, our application of the prioritization tool is a deliberative, pre-decisional process and represents only one of multiple factors informing agency decision-making. The scores and rankings generated from the prioritization tool are pre-decisional and are subject to change as new information becomes available. Additionally, stakeholders have indicated that published scores could create negative perceptions about substances simply because they were evaluated using the tool. As such, we do not plan to publish prioritization scores or rankings, but we will announce chemicals selected for full scientific assessment and will list these chemicals and the status of their reviews on our dedicated webpage, *List of Select Chemicals in the Food Supply Under FDA Review*.[3] Additional information about transparency in the post-market systematic review process is detailed in the companion paper.

The pilot of version of the tool (Version 1) was tested internally, involving scoring of a set of chemicals and performing sensitivity analysis. We received public comment[4] and feedback on Version 1, following FDA's public meeting entitled "Development of an Enhanced Systematic Process for FDA's Post-Market Assessment of Chemicals in Food" held in September 2024.[5] We then updated the tool (Version 2) to include the toxicity rubric to score the 'Toxicity' criterion. Version 2 of the tool underwent additional public comment[6] in 2025 and this external peer review. This external peer review, the public comment, and additional internal testing informed the current version of the tool (Version 3), in which we have added additional detail regarding the systematic post-market assessment process.

We agree with the peer reviewers that unintentional contaminants and intentionally added substances require distinct consideration and scientific assessment. The prioritization tool is designed to accommodate both chemical classes and will be applied separately by subject matter experts in different regulatory programs within the Human Foods Program.

We have addressed the more topic-specific comments under the summary/response sections relevant to those respective topics.

[3] https://www.fda.gov/food/food-chemical-safety/list-select-chemicals-food-supply-under-fda-review
[4] https://www.regulations.gov/document/FDA-2024-N-3609-0001
[5] https://www.fda.gov/food/workshops-meetings-webinars-food-and-dietary-supplements/public-meeting-development-enhanced-systematic-process-fdas-post-market-assessment-chemicals-food
[6] https://www.regulations.gov/docket/FDA-2025-N-1733

Question 2: The draft Post-market Assessment Prioritization Tool currently includes four Public Health criteria and three Other Decisional criteria.

2.a. Are the four Public Health criteria appropriate for the purpose of the tool? If not, please explain what changes might be considered and why.

There was consensus among the reviewers that the four Public Health criteria are appropriate. One reviewer supported the types of information included in the four Public Health criteria but suggested an alternative to the current approach be explored; specifically, the reviewer discussed the use of a 'Risk criterion,' which would, as suggested here and elsewhere in this document, incorporate the concepts of exposure and toxicity into a single value, such as a point of departure (POD), acceptable daily intake (ADI), or similar. Of note, Threshold of Toxicological Concern (TTC), Expanded Decision Tree (EDT), or similar values are also metrics that incorporate exposure and toxicity and have been suggested here and elsewhere in this document by reviewers.

There were specific comments/suggestions regarding the scoring definitions and weighting of the Public Health criteria, which are also addressed in detail in Q3.a. and Q4.a.

Reviewers had concerns with the definition of the 'Change in exposure' criterion. Two reviewers asked for more information regarding how the magnitude of the change in exposure relates to the exposure level, the direction of change, and whether a change approaches or exceeds the TTC. One reviewer suggested FDA consider whether chemicals not previously assessed should have a higher score for 'Change in exposure.' One reviewer said it would be helpful for FDA to provide rationale behind the 'Change in exposure' criterion rather than defining an exposure criterion based on current level of exposure in the food supply, which the reviewer says is the most appropriate input to a risk-based score. Utilizing evidence from human biomonitoring studies was also suggested as a potential source of information for scoring the 'Change in exposure' criterion.

One reviewer raised concerns about redundancy or double counting, including
(i) Whether scoring for 'Susceptible subpopulation' may result in double-counting if there were a prior assessment that already considered a vulnerable subpopulation;
(ii) Potential double counting of 'New scientific information and potential impact' criterion with the other three Public Health criteria for a previously assessed chemical.

Reviewers asked that the term 'Susceptible subpopulation' be clearly defined to avoid any ambiguity and ensure consistency. One reviewer asked whether the 'Susceptible subpopulation' criterion is a general assumption or chemical specific.

Regarding the 'New scientific information and potential impact' criterion, one reviewer commented that a weight of evidence approach must support the new information before it can justify prioritizing a chemical for reassessment.

FDA Response:

We appreciate the detailed comments from the reviewers. The four Public Health criteria were retained and the scoring definitions for each were revised. We address the suggestions to add a 'Risk criterion' in Q2.c.

FDA thanks the reviewers for carefully considering the 'Change in exposure' criterion and especially how it relates to the current level of exposure in the food supply. As noted, we revised the scoring definitions to provide clarity. For authorized food additives, color additives, and GRAS substances, the current level of exposure from the food supply should reflect an exposure of no more than the currently authorized and GRAS intended uses for the substance. These intended uses are specific and result in a defined dietary exposure profile. The existing authorizations and intended uses determine this exposure profile, which we refer to as the estimated daily intake (EDI) or cumulative estimated daily intake when all sources of dietary exposure are considered together. The level of exposure that is considered safe, or in other words, that is associated with a reasonable certainty of no harm[7] is referred to as the ADI. This ADI is determined based on all available toxicology, epidemiology, safety, and history of use information for the substance, and in some cases, related substances. The key reason why 'Change in exposure' is a critical criterion, rather than simply exposure itself, is because the current set of intended uses for most of the food chemicals FDA has previously assessed are designed to keep the EDI safely at or below the ADI level. Any factors or new knowledge that indicate a potential *increase* in exposure should trigger greater concern for prioritization because these factors may result in the EDI exceeding the ADI or EDT/TTC value (where applicable), thus no longer providing reasonable certainty of no harm. In the case of contaminants, changes in exposure could indicate a new or previously unrecognized source of contamination (e.g., PFOA leaching from high-density polyethylene pesticide containers),[8] warranting a higher score in the prioritization process. Additionally, we note that the 'Change in exposure' criterion is sensitive to directionality of the change – if there is no change or a decrease, a lower score is returned, while an increase returns higher scores, and a change of unclear direction returns a moderate score. This scoring function reflects the level of concern for the direction of change and its potential impact on the ADI to EDI comparison.

As suggested by one reviewer, we are currently exploring methods to utilize human biomonitoring data into our dietary exposure assessments. However, since this would likely be a time-intensive process, we do not expect to utilize this type of data during prioritization. Rather, it could be useful during the full scientific assessment of a food chemical.

Regarding the concerns of whether the 'Susceptible subpopulation' criterion was chemical specific or general and how this may/may not impact inflation of this score (double-counting), we appreciate the reviewer's questions and have made clarifications

[7] See 21 CFR 170.3(i)
[8] EPA Releases Data on Leaching of PFAS in Fluorinated Packaging | US EPA

to address this concern. As also discussed in Q3.a.i., we revised the criterion definition and updated the name of this criterion ('Use or presence in food for susceptible subpopulation') to more clearly reflect that it is the vulnerability of the intended consumers of the food containing the chemical (intended for specific populations that may have vulnerabilities) and not the chemical itself, that is being considered under this criterion. For example, if a food chemical undergoing prioritization is found to be highly utilized in breakfast cereals marketed to children and, given that children may have unique windows of toxicological susceptibility (i.e., they are a susceptible subpopulation in this context), we would want to ensure that the final prioritization score captures this increased concern.

There were two double-counting concerns raised by a reviewer. The first concern related to a potential double-counting of the 'Susceptible subpopulation' criterion if a susceptible subpopulation had been considered in a previous assessment of a food chemical. The scoring of the 'Use or presence in food for susceptible subpopulation' criterion is independent of whether the chemical was previously assessed or reviewed by FDA. The intention of this criterion is to elevate the prioritization scores of food chemicals that are present in food specifically marketed to or intended for susceptible subpopulations. To provide clarity, as we discuss above and in Q3.a.i., we renamed the criterion to 'Use or presence in food for susceptible subpopulation' to indicate that the food(s) in which the food chemical is found is the subject of this criterion and not simply the population. The second concern was related to a potential double-counting of the 'New scientific information and potential impact' criterion with the other three Public Health criteria for food chemicals that have been previously assessed. We acknowledge there is potential for overlap in the scoring of food chemicals for the 'New scientific information and potential impact' criterion and other criteria; for example, a new toxicity study may impact the score assigned for Toxicity and also the score for 'New scientific information and potential impact.' As we discuss in Q3.a.i, the scoring definitions for the 'Change in exposure' criterion and 'New scientific information and potential impact' criterion were revised, including revisions for scoring of chemicals not previously assessed or reviewed by FDA. We consider any remaining potential double-counting to be acceptable in this instance because new information for a food chemical is exactly the type of data that could change the conclusions of previous assessments. One of the core goals of the prioritization tool, especially when applied to intentionally added food chemicals, is to elevate the prioritization scores of food chemicals that have new scientific information that might, upon full scientific assessment, result in the EDI exceeding the ADI.

As discussed in response to Q3.a.i., we have revised the scoring definition for this criterion to more clearly address chemicals that were not previously assessed.

We agree with the reviewer who suggested a weight of evidence approach be used; a weight of evidence approach will be used for scoring of the 'Toxicity' criterion using the toxicity rubric as well as during the assessment stage of the post-market systematic process if the food chemical is prioritized for a full scientific assessment. Regarding the 'New scientific information and potential impact' criterion, we have specified in the

definition that our goal is to identify data and information that "would impact or change the conclusions of the previous assessment or review." Thus, all new relevant scientific information will be screened to determine its impact on FDA's previous scientific conclusions.

2.b. Are the three Other Decisional criteria appropriate for the purpose of the tool? If not, please explain what changes might be considered and why.

Two reviewers suggested that FDA consider removing the Other Decisional criteria altogether. One reviewer raised concern about double counting Other Decisional criteria and the Signal Monitoring of FDA's Systematic Process; this reviewer suggested that the Other Decisional criteria only be part of Signal Monitoring process. One reviewer commented that the Other Decisional criteria are important but are not appropriate for inclusion in the tool's primary decision framework.

There were concerns about the Other Decisional criteria being highly subjective, especially the 'Building public confidence' and 'External activity/attention' criteria. One reviewer supported including 'Other governmental decisions' criterion in the tool, while another reviewer recommended the 'Other governmental decisions' criterion be weighted more heavily after further defining to focus on comprehensive science-based risk assessments by other regulatory authorities and scientific bodies that could impact FDA's previous safety assessments. Another reviewer commented that the Other Decisional criteria are appropriate but that there may be challenges to quantitatively combining Other Decisional criteria with total Public Health criteria.

One reviewer commented that high attention may come and go quickly and suggested removing the 'External activity/attention' criterion. The reviewer commented that external activity/attention should instead be addressed outside the tool via rapid response capability and ongoing communication and outreach activities.

Concerns with the 'Building public confidence' criterion were raised and reviewers recommended removing that criterion.

One reviewer commented that two separate lists for intentionally added substances and unintentional contaminants could be useful and facilitate communication.

FDA Response:

We appreciate the comments/suggestions from the reviewers regarding the Other Decisional criteria. We have removed the Other Decisional criteria from the prioritization tool altogether. The justification for doing so, in part, is that the post-market systematic process has been further defined and includes an internal triage process (occurring before prioritization) that captures aspects of the Other Decisional criteria. This revision addresses the many concerns raised by reviewers. Overall, we agree with reviewers that Other Decisional criteria are important but may be better suited outside of the prioritization tool. We added more information to the Overview section in the revised tool description document, and the companion paper to the

revised tool is being published which includes more details about triage and how the Post-market Assessment Prioritization Tool fits into the broader systematic post-market assessment process.

We have addressed the comment regarding separate lists for intentionally added substances versus unintentional contaminants in our response to Q1.

2.c. Are there additional criteria that should be considered? If so, please describe additional criteria that might be considered and why.

Two reviewers commented that no additional criteria should be considered.

Three reviewers suggested additional or alternative Public Health criteria for consideration. One reviewer suggested a quantitative risk-based criterion to understand the likelihood that a chemical poses a risk and suggested this could be assessed by using Predicted Exposure/Estimated Acceptable Daily Intake. One reviewer suggested a Risk criterion, as discussed in Q2.a. One reviewer suggested adding an aggregate exposure criterion. Another suggestion was to add a criterion to indicate whether a chemical is part of a family of chemicals of similar toxicity.

One reviewer suggested an additional Other Decisional criterion for consideration: value to the food or the food supply.

FDA Response:

We thank the reviewers for their comments about additional criteria for consideration.

This prioritization tool includes risk factors (e.g., toxicity). After prioritization, if a substance is identified as a high priority chemical, the subject matter experts can proceed with a comprehensive risk assessment. By leaving the exposure and hazard components as separate independent elements, the prioritization tool has greater potential to distinguish among substances and retain information indicating the main drivers of the score rather than folding this information into a single risk value (e.g., POD, ADI, TTC value). Regarding the other suggested Public Health criteria, consideration of cumulative dietary exposures and chemically/ pharmacologically similar substances are indeed important factors which are addressed during the comprehensive assessment step of the post-market systematic process (i.e., after prioritization).

We appreciated the suggested additional Other Decisional criteria; however, as discussed, we have removed the Other Decisional criteria from the prioritization tool.

Question 3: The draft scoring definitions for all criteria were developed to consider the expected variability in the types and extent of data available for the wide variety of food chemicals that may be considered for review.

3.a. Given this context, are the scoring definitions for the Public Health criteria appropriate for the purpose of the tool?

One reviewer recommended expanding the scoring range of Other Decisional criteria to 1-12 instead of 1-9, with the intention that greater spacing between score levels will create greater variability in the combined score and should help distinguish the ranked chemicals as needed.

One reviewer noted that the prioritization tool's focus on the traditional toxicological approach and previously assessed substances makes sense for post-market surveillance but results in lower overall scores for all contaminants and non-intentionally added substances lacking a previous assessment by FDA.

One reviewer commented that the 'Toxicity' criterion should be based on a weight of evidence approach. We addressed this in response to comments in Q3.a.ii. and Q5.

FDA Response:

We thank the reviewers for their suggestions to improve the scoring definitions for the Public Health criteria.

We agree that sufficient scoring differentiation is important for prioritization. Changing the criteria scoring range from 1-9 to 1-12 would indeed increase the number of unique overall scores (i.e., scoring differentiation). However, based on our testing, the scoring range of 1-9 currently selected is sufficient and consistent with standard approaches. However, in the future we could consider changing the scoring functions (e.g., increasing the range of scores while keeping the range the same for all criteria as suggested by the reviewer).

Regarding potential differences in scoring for different types of food chemicals and the status of previous assessment by FDA, the tool has been designed for the post-market prioritization of chemicals in food, including food additives; color additives; generally recognized as safe substances; food contact substances; and chemicals that are present as unintentional contaminants. Based on our internal testing of the tool, we do not see systematically lower overall scores for chemicals that are present as unintentional contaminants.

3.a.i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why?

Reviewers provided several specific comments/suggestions regarding each of the four Public Health criteria scoring definitions. For 'Toxicity' and 'Change in exposure,' reviewers provided similar comments in Q3.a.i. that we've addressed in responses to comments provided in Q2.a.

Two reviewers noted there is some ambiguity in the scoring definitions that require expert judgement. One reviewer suggested clarifications for 'Change in exposure' scoring definitions and suggested renaming the criterion to 'Change in exposure and potential impact.'

Two reviewers suggested we should more clearly define 'Susceptible subpopulation,' and one reviewer raised concern that the criterion had broad applicability (i.e., most chemicals can affect some subpopulation).

Regarding the 'New scientific information and potential impact' criterion: One reviewer commented that the definitions are not applicable for substances not previously assessed by FDA and asked what the scoring would be for those substances. Another reviewer suggested grouping potential low impact information with no new information or expanding the scoring so that information with an uncertain impact would have its own score.

FDA Response:

We thank the reviewers for comments and suggested revisions.

We revised the 'Change in exposure' scoring definitions to provide clarity, with discussion of this criterion also found in Q2.a. Rather than revising the criterion name, as suggested by the reviewer, we retained the original criterion name; potential impact is assessed during the full scientific review, which would occur after a substance has been prioritized.

We also revised the 'Susceptible subpopulation' definition to provide clarity. In response to the reviewer's concern that most chemicals affect some subpopulation—while this may be true, the criterion is defined by a chemical being found or potentially present in food specifically marketed to or intended for susceptible subpopulations; we revised the definition and included examples. We renamed the criterion to be 'Use or presence in food for susceptible subpopulation' so that it is clearer. Additionally, we revised the scoring definition for score = 3 to be "Unable to assess due to insufficient information" for clarity and consistency with other criteria.

We revised 'New scientific information and potential impact' to clarify how to score a substance that was not previously assessed or reviewed by FDA. We agree with and have implemented the suggested revisions to the scoring definitions for scores of 1, 3, and 5, respectively, as detailed in the revised tool description document.

3.a.ii. The toxicity criterion described in Section 3.1.1 considers data for seven different toxicity data types and the score assigned reflects the highest toxicity data type score from the toxicity rubric, which is described in Appendix A Table A1. Is this the most appropriate strategy for assigning a toxicity criterion score? If not, please explain your reasoning and provide alternatives for FDA to consider. Please be specific and provide references, as appropriate.

There was a consensus among reviewers that assigning the highest toxicity score is appropriate or adequate for the purpose. One reviewer asked for clarification as to whether FDA intends to always address all endpoints and information sources described in Table A1 (i.e., the toxicity rubric) and demonstrated in Table A2, or whether it is sufficient to identify the single critical toxic effect which results in that substance receiving the highest score from the rubric (i.e., further toxicity rubric scoring would cease and the high score would forward into the main tool).

One reviewer raised concern that substances without data could be given an advantage (we understood this to mean a lower toxicity score) over substances with some data and recommended FDA consider splitting the intermediate score of 5 into: score = 5: no data and no evidence for toxicity, score = 7: no data and evidence for toxicity from quantitative structure activity relationship (QSAR) models. One reviewer raised concern that the focus on qualitative attributes may lead to misleading scoring and instead suggested a prediction of a POD or ADI as an output of the 'Toxicity' criterion. Related, in response to this question and others, at least one reviewer suggested consideration of TTC (or EDT, or similar) as a possible means of bridging data gaps for understudied candidate substances. Reviewers also noted that these tools (e.g., TTC) are intended to be protective for genotoxic and other carcinogens, except in cases where substances are in "chemical cohorts of concern."

It was also suggested that FDA consider whether the acute toxicity scoring parameter was necessary and noted that the term "activity" as used in the toxicity rubric be explained further or removed.

It was suggested that FDA make further use of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) and other authoritative body designations where possible, including potentially also using this approach for developmental and reproductive toxicity or endocrine criteria. Related to this point, one reviewer recommended refinement of how carcinogen and/or genotoxicity signals are handled by the toxicity rubric, suggesting priority be given to capturing genotoxic substances either through use of classifications first (e.g., International Agency for Research on Cancer, GHS), and then through a weight of evidence approach (e.g., a battery of genotoxicity studies such as those from the Organisation for Economic Co-operation and Development test guidelines) if classifications have not been established. The rationale provided by the reviewer was that carcinogens (presumably oral) should not be authorized as food additives. Another reviewer expressed concern that the human health risk from different carcinogens may vary depending on the mode of action (e.g., genotoxic vs non-genotoxic) and that the toxicity rubric may not reflect this underlying scientific knowledge. Additionally, a reviewer inquired about the triage process and asked if genotoxic carcinogens might bypass the prioritization tool and proceed directly to risk assessment.

One reviewer suggested here and elsewhere that a Risk criterion be incorporated into the Public Health criteria, of which, the toxicity criteria would be a foundational aspect. If such a Risk criterion is introduced, the reviewer suggested that it would then be possible to move the bioaccumulation/biopersistence criterion to a stand-alone criterion within the larger set of Public Health criteria.

One reviewer noted the heavy reliance on animal toxicity data within the toxicity rubric and noted that the concordance of toxicological endpoints between clinical and nonclinical data is known to be poor.

Reviewers also provided specific comments and requests for clarification about Table A1.

FDA Response:

We thank the reviewers for the comments/suggestions regarding scoring of the 'Toxicity' criterion. Yes, FDA will always score all the endpoints in the toxicity rubric rather than stop when at least one parameter has reached the maximal score (which advances back to the main tool). We have since prepared a revised tool description document to ensure this is explicitly stated, including adding a footnote in Table A2 to note that all endpoints will be scored. The purpose of this is to retain as much information as possible and document which signals were considered and available at the time of scoring. It also generates a working file with more complete information in the case where a chemical does not advance to prioritization initially but does undergo scoring in subsequent rounds of tool utilization.

Regarding the concern about scoring of substances with and without available toxicity data, and in response to utilizing the EDT and related tools (Q5), we have updated the toxicity rubric to leverage EDT class assignments in cases where data are insufficient to provide a weight of evidence evaluation for the target organs or endpoints within the toxicity rubric. Additional details are provided in response to Q5. This update also addresses one reviewer's suggestion that a score of 7 be added to the scoring options to split the current score of 5 when data are unavailable. We anticipate that this will make it unnecessary to split the 5 score into a 5 and a 7. However, this recommendation could be revisited in future versions of the prioritization tool should we find it necessary. We also note that during our most recent trials of the toxicity rubric conducted as part of sensitivity analysis, we found that substances lacking toxicology or other safety data had somewhat inflated scores compared to substances with safety information. This was indicated by multiple data-poor food chemicals receiving scores of 5 while those chemicals with larger safety databases had more evenly distributed scores across the toxicity rubric criteria. As such, we implemented an 8th scoring option for safety database compensation. Please refer to Q5.b. for more information on the scoring of data-poor food chemicals.

In response to the suggestion to utilize a prediction of POD or ADI in lieu of qualitative attributes for scoring the toxicity rubric, we have decided not to include a POD or ADI estimation at this time. Since we cannot ensure that all substances undergoing prioritization will have suitable safety databases from which to estimate a POD (e.g., the no observed adverse effect level may be the highest dose tested for naturally-derived substances), we prefer to maintain the current level of flexibility for capturing qualitative toxicity aspects (signals). Also, the tool must be capable of capturing emerging signals from data streams other than conventional toxicology studies, and these signals may not be ideal for POD or ADI derivation. An example of such a signal would be increased risk of heart disease with consumption of artificial *trans* fat.

We respond to the reviewer's suggestions regarding a Risk criterion in our response to Q.2.c.

We agree with the suggestion that food chemicals with new evidence of carcinogenicity should bypass the prioritization process. Such new evidence might include, but is not limited to, a new animal or epidemiology study or studies, assessment or reevaluation of existing studies or data, or new information about the carcinogenicity of similar substances. We have incorporated this into the triage step of the post-market assessment process.

We have revised Table A1 to address requests for clarification and related comments. Importantly, we have clarified that a weight of evidence approach should be taken for each toxicological endpoint scored in the toxicity rubric.

3.b. Are the scoring definitions for the Other Decisional criteria appropriate for the purpose of the tool?

As mentioned earlier, two reviewers strongly recommended removing the Other Decisional criteria from the tool altogether. Other reviewers recommended removing the 'Building public confidence' criterion.

One reviewer suggested expanding the scoring range of Other Decisional criteria to 1-12 instead of 1-9, as also suggested for the Public Health criteria.

FDA Response:

As discussed, the Other Decisional criteria have been removed from the tool.

We responded to the reviewer's suggestion about the scoring range in Q3.a.i.

3.b.i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why.

Three reviewers commented that the definitions are appropriately defined. Two reviewers suggested revisions for the 'External stakeholder activity/attention' criterion; one reviewer recommended revising the definition for score 3 to be "conflicting" rather than "uncertain;" and another reviewer commented that clarification between score of 9 and 5 would be helpful. One reviewer asked for clarification as to what is meant by 'Building public confidence' and commented that it may be especially challenging to assign a score given the current definitions. One reviewer commented that the Other Decisional criteria can be strongly influenced by unrelated events occurring at the same time that may receive more attention (e.g., reporting in the news/social media).

FDA Response:

We thank the reviewers for suggested revisions and agree with the need to provide clarification. However, as discussed in Q2.b., the Other Decisional criteria has been removed from the tool.

3.b.ii. FDA is exploring quantitative and qualitative methods to help inform the scoring of the 'building public confidence' criterion (Section 3.2.3) such as conducting public

sentiment analysis (e.g., utilizing natural language processing). How might such tools or the information they provide be incorporated into this criterion? What additional strategies and metrics could FDA consider?

Three reviewers recommended that the 'Building public confidence' criterion should not be included in the tool. Two reviewers discussed the importance of FDA building and maintaining public confidence but recommended that measurement of this criterion be conducted outside of the tool, with one reviewer stating that assessing public confidence should be an ongoing or at least periodic activity by the Agency. Another reviewer emphasized that transparency and communication are critical for public confidence. One reviewer asked how 'Building public confidence' will be measured.

Two reviewers raised concerns about the data used for public sentiment analysis. There was a concern of whether there would be public sentiment data on individual ingredients or food additives beyond those ingredients/additives already widely known. One reviewer concurred with the public comments that public sentiment data gathered by applying natural language processing on social media posts could be "gamed" through use of technology or social media campaigns and thus bias the scoring. Another reviewer commented that analysis informing this criterion should not be limited to text data from the internet.

FDA Response:

As discussed above, we removed the Other Decisional criteria from the tool. We agree that building and maintaining public confidence is important. We appreciate the suggestion to assess public confidence as a periodic activity, and FDA will consider establishing an approach to evaluate public confidence that is measurable.

We thank the peer reviewers and public commentors for bringing to our attention concerns about data used for public sentiment analysis and possible biases.

Question 4: The prioritization methodology includes weighting factors. FDA is considering equal weighting among the Public Health criteria and (separately), among the Other Decisional criteria for the Post-market Assessment Prioritization Tool.

One reviewer commented that unequal weighting may or should be considered. One reviewer recommended that greater weight be given to Public Health criteria over Other Decisional criteria. Another reviewer commented that weights should reflect the relative importance of each criterion to the FDA decision making process and if equal weights are to be used then the reasoning for this decision should be provided.

FDA Response:

We thank the reviewers for considering this question and note that since we have removed Other Decisional criteria in response to reviewers' suggestions, this question is no longer relevant.

4.a.i. Should different weights be applied to the Public Health criteria when determining the Total Public Health Criteria Score? If so, please specify the weighting scheme that might be considered and why.

Two reviewers commented that equal weighting of the Public Health criteria is reasonable. Another reviewer commented that FDA should provide a clear justification for the weighting scheme that is selected considering the relative importance of the criteria.

Some reviewers recommended an alternative weighting scheme be considered. One reviewer suggested increasing the weight of 'Toxicity' and decreasing the weight of 'Susceptible subpopulation.' One reviewer commented that it was challenging to address this question because several of the Public Health criteria require more explicit definition, but that as currently described (i.e., prior to revisions for clarity), relatively more weight should be given to 'Toxicity' and 'New scientific information and potential impact' criteria. One reviewer suggested as an alternative weighting scheme that 'Change in exposure,' 'Susceptible subpopulation,' and 'New scientific information and potential impact' modulate a main criterion such as a Risk criterion (see Q2.a., Q2.c., and Q3.a.ii.) with an example starting point of equal weighting of the main criterion and the sum of the remaining criteria (i.e., weight of 0.5 for the main and weights of 0.5/3 for the other three criteria). One reviewer discussed uncertainty, modulators, and the challenges of weight selection and provided suggestions to consider for this or a future version of the prioritization tool.

One reviewer noted that the 'Toxicity' rubric covers acute effects and that these effects should be weighted differently compared to long-term effects and recommended considering severity and exposure level in this or a future version of the prioritization tool. This was specifically suggested in the context of a Risk criterion.

FDA Response:

We thank the reviewers for the comments/suggestions regarding weighting of the Public Health criteria. We agree that it is challenging to select a weighting scheme and that it is

important to provide reasoning for the selected scheme. Overall, there was no consensus among the reviewers about an alternative weighting scheme. Multiple reviewers suggested increasing the weight of the 'Toxicity' criterion; however, we do not see a robust scientific rationale for only increasing this criterion weight. Sensitivity analysis was conducted on the revised tool to include the following scenarios: equal weighting as baseline; increasing weight for 'New scientific information and potential impact' and 'Change in exposure;' and increasing weight for 'Toxicity' because of the reviewers' suggestions. While there were small shifts in the overall scores, substances with higher overall scores remained high and substances with lower overall scores remained low. Based on our testing, scoring differentiation using equal weighting for determining the overall score is currently sufficient. Thus, equal weighting of the Public Health criteria is used for determining the overall Post-market Assessment Prioritization Score. However, in the future we could consider changing the weighting scheme if needed (and as the FDA Human Foods Program deems appropriate) as acknowledged in the revised tool description document.

We appreciate the reviewer's consideration that alternative weighting might be considered for acute effects compared to chronic effects or longer-term effects in combination with a Risk criterion. In the FDA Human Foods Program, systems separate from the prioritization tool and systematic post-market review process are in place to capture acute adverse event signals from substances in food. Once these acute adverse events are identified, relevant safety and regulatory processes are implemented to protect the public health. At this time, we do not plan to make changes to the acute category of the toxicity rubric. As incorporated into the rubric, the acute scoring category will serve its original purpose, which is to capture a measure of the inherent toxicity of substances being scored. The consideration of a Risk criterion is discussed in Q2.c.

4.a.ii. Should different weights be applied to the Other Decisional Criteria when determining the Total Other Decisional Criteria Score? If so, please specify the weighting scheme that might be considered and why.

Two reviewers suggested removing the Other Decisional criteria from the tool altogether. One reviewer commented that equal weighting is appropriate, while another reviewer commented that equal weighting is a reasonable starting point. One reviewer suggested removing the 'Building public confidence' criterion and weighting 'Other governmental decisions' 80-100% (i.e., remove the other two Other Decisional criteria or give the most weight to this criterion and partial weight to 'External activity/attention').

FDA Response:

We thank the reviewers for the comments/suggestions regarding Other Decisional criteria weighting. As discussed, the Other Decisional criteria have been removed from the tool.

4.b. FDA is considering equal weighting among the Total Public Health Criteria Score and the Total Other Decisional Criteria Score to determine the overall Post-market Assessment Prioritization Score.

Reviewers provided suggestions including removing the Other Decisional criteria from the tool altogether, weighting Public Health criteria score more heavily when calculating the overall score, or using equal weighting. More detailed (and related) comments were provided in Q4.b.i.

FDA Response:

We thank the reviewers for the comments/suggestions regarding the respective weighting of the Total Public Health Criteria Score and the Total Other Decisional Criteria Score for determination of the overall Post-market Assessment Prioritization Score. As discussed, the Other Decisional criteria have been removed from the tool.

4.b.i. Should different weights be applied when determining the overall Post-market Assessment Prioritization Score? If so, please specify the weighing scheme that might be considered and explain why it would be more appropriate than equal weighting.

Two reviewers suggested removing the Other Decisional criteria from the tool altogether. One reviewer recommended greater weight for the Total Public Health Criteria Score and suggested two options for weighting schemes.

Two reviewers raised concern about score differentiation (i.e., that many substances may receive the same overall score) and suggested alternative weighting and other approaches for determining the overall Post-market Assessment Prioritization Score. One of the suggestions was to remove the Other Decisional criteria from the tool and instead determine the overall score solely based on Public Health criteria and assign a value (e.g., a, b, c) after from the Other Decisional criteria. One reviewer suggested three separate prioritized lists—one for the Public Health criteria, one for the Other decisional criteria, and one for the overall score.

Two reviewers commented that FDA should provide justification for the selected weighting scheme.

FDA Response:

We appreciate the suggestions for different weighting schemes and approaches to determining the overall Post-market Assessment Prioritization Score. As discussed, the Other Decisional criteria have been removed from the tool, and as such the overall Post-market Assessment Prioritization Score will be determined based on Public Health criteria only. The weighting of Public Health criteria is addressed in Q4.a.i. and in the revised tool description document.

Question 5. The draft toxicity rubric uses traditional toxicity data (in vivo, as well as limited in vitro such as for genotoxicity), human health outcomes (e.g., adverse event reports), and epidemiological data for determination of the toxicity criterion score within the Public Health criteria. Considering that the prioritization process is not a comprehensive review, please address the following questions.

5.a. How might FDA incorporate information from new approach methodologies (NAMs) into the toxicity rubric?

Reviewers noted, in response to this question and elsewhere, that NAMs are currently incorporated into some aspects of the toxicity rubric (e.g., in the weight of evidence evaluation of genotoxicity). There were suggestions that this could be further formalized, for example, by stating directly that a conventional genotoxicity testing battery could be utilized in the genotoxicity evaluation component.

Beyond this, reviewers did not cite specific instances where NAMs could be leveraged, but they did provide some broad thoughts on the topic. Reviewers noted the distinction between the use of NAMs for mechanistic insight and hazard identification and the challenge of their adoption for regulatory use but noted that *in silico* approaches may have value in terms of toxicity predictions or where data are lacking (e.g., QSAR, EDT).

One reviewer also noted that a flexible framework will allow gradual improvement in NAMs integration as new methods advance and become more established (e.g., artificial intelligence).

FDA Response:

We thank the reviewers for considering how NAMs might be incorporated into the toxicity rubric. In response to reviewer comments to this question and elsewhere, FDA has clarified and expanded the language in the toxicity rubric related to evaluating the weight of evidence for genotoxicity and/or carcinogenicity. We now clearly reference the concept of using a genotoxicity battery of testing where available, as well as noting the incorporation of other signals relevant to assessing the potential for genotoxicity or carcinogenicity.

5.a.i. Are there specific NAMs (e.g., systems biology, engineered tissues, artificial intelligence, in vitro, microphysiological systems, or other alternative data or modeling tools) that would be most appropriate for use in the toxicity rubric? If so, please explain which NAM(s) would be most appropriate and why.

Reviewers noted that a major advantage of certain NAMs is the use of human cells and tissues which may have the potential to more accurately determine chemical toxicity. However, reviewers also noted the lack of xenobiotic metabolism as a significant limitation of many NAMs. While reviewers did not explicitly recommend specific NAMs for inclusion in the toxicity rubric, in combination with responses to the full set of questions within Q5 and elsewhere, reviewers generally supported NAMs as one component of many in an overall weight of the evidence approach to the toxicity rubric.

FDA Response:

We thank the reviewers for considering possible NAMs to incorporate into the tool. As recommended by the reviewers, FDA continues to consider NAMs as part of the weight of evidence approach for the toxicity categories outlined in the rubric. In response to this (Q5) and other peer review comments, we have clarified the notion of "activity" expressed in the toxicity rubric by changing this word to "signal." We hope that this better reflects our intention to capture signals from NAMs based evidence where available to support the toxicity scoring.

5.a.ii. Given that a single NAM is not expected to be a one-to-one replacement for a traditional in vivo toxicity test, how can the strengths and limitations of each NAM be appropriately considered if it is incorporated into the toxicity rubric?

Specific NAMs were not recommended by the reviewers for incorporation into the toxicity rubric in response to the previous question. Specifically, reviewers cautioned that NAMs should not be considered as replacement for conventional animal toxicity information, and stressed awareness around issues related to the lack of xenobiotic metabolism, extrapolation of results to human relevant endpoints, and preference for validated NAMs. Reviewers also noted that NAMs or tools such as EDT could be leveraged to derive toxicological benchmarks, but that NAMs also have utility in the hazard identification phase and should be considered as part of the weight of evidence during the toxicity rubric scoring.

FDA Response:

We thank the reviewers for providing a few points for consideration in response to this question that relate to potential strengths and limitations should a NAM be leveraged during evaluation of the toxicity rubric. As recommended by the reviewers, FDA considers NAMs as part of the weight of evidence for the specific toxicity parameters included in Table A1, and we have included additional notes to ensure that users of the tool consider these signals for each indicated category of information in the rubric. FDA acknowledges the challenges cited by the reviewers (metabolism, extrapolation, validation) and considers these and similar factors when reviewing NAMs in weight of evidence evaluations using the toxicity rubric. FDA concurs with reviewers that NAMs would not replace conventional toxicity data but rather be used to compliment it during evaluations and scoring.

5.b. Threshold of Toxicological Concern (TTC) approaches can be used to assess the toxicity of chemicals that lack sufficient safety data and have low dietary exposures. Although the Cramer classification scheme has historically been used in TTC approaches, FDA has recently developed the Expanded Decision Tree (EDT) that assigns chemicals to one of six EDT classes. How might such tools or the information they provide be incorporated into the toxicity rubric?

In general, reviewers supported the use of the EDT, TTC, or similar approaches, in cases where there is limited toxicity information for a candidate substance, but reviewers also highlighted the

challenge that an automated version of EDT is in development, but not yet fully available for use. Consideration of EDT, TTC, or similar was suggested by reviewers for providing exposure thresholds of concern in situations where genotoxicity/ carcinogenicity information is lacking. Additionally, one reviewer noted the challenges in potentially utilizing the EDT or TTC for scoring the "other organ-specific toxicity" category. Another reviewer suggested that if there is inadequate safety data, the EDT could be substituted to provide an overall score for the toxicity rubric.

Reviewers also mentioned the concept of using the EDT or similar if a more risk-based approach is taken to the overall tool rather than the hazard-based approach currently used in the toxicity rubric. Specifically, EDT (as well as TTC) incorporates an exposure component ("how high would exposure need to be for there to be a concern?") in conjunction with the hazard component. In contrast, the current format of the toxicity rubric is structured to consider hazard only, with exposure considerations being incorporated as a separate criterion in the Public Health criteria, rather than in the toxicity rubric.

FDA Response:

We thank the reviewers for considering the potential use of the EDT within the toxicity rubric of the prioritization tool. FDA concurs that, in the absence of information regarding genotoxicity/ carcinogenicity, a tool such as EDT or TTC could be useful in establishing exposure-based thresholds of concern. At present, a lack of such information would result in score of "Moderate (5)" for this parameter which, depending on other factors, may or may not result in the candidate substance being prioritized for full assessment.

Scoring trials using the toxicity rubric during sensitivity analysis suggested that the scores of data-poor food chemicals may be inflated compared to those with more data; this was indicated by multiple data-poor food chemicals receiving scores of 5, while those with larger safety databases had more evenly distributed scores across the toxicity rubric criteria. In response to reviewer suggestions and this additional information, we have incorporated EDT within the toxicity rubric for cases of data-poor substances. However, because the toxicity rubric is intended to decouple inherent toxicity characteristics from exposure (exposure is investigated and scored separately in the 'Change in exposure' criterion), we utilized only the EDT class assignment. Data-poor substances will still be scored to the extent possible using the toxicity rubric, but they will also receive a database compensation score that captures the level of concern for the substance based on its structure. Substances that are placed in EDT Class I or II will receive a database compensation category score of 1 in the toxicity rubric, while substances placed in EDT Class III or IV will receive a 5, and substances placed in EDT Class V or VI will receive a 9. We plan to review this process in the future to determine whether adjustments may be needed.

As the larger post-market review process is presently designed, risk (as a product of hazard and exposure) will be fully considered once a substance has been prioritized for assessment. At that point, EDT or TTC would be leveraged where applicable, to compliment any existing data and fill data gaps.

Discussion of the risk-based criterion suggested in some of the reviewer comments is included in response to Q2.c.

Question 6. Do you have any additional comments? Please share them in your review.

There were some additional suggestions/comments. Only those comments that were not addressed in previous questions are discussed here.

One reviewer suggested considering scoring standardization such that scoring be e.g., {1, 5, 9} for each criterion instead of {1, 5, 9} or {1, 3, 5, 9}.

Two reviewers commented about the importance of transparency. One reviewer commented about the need for transparency of the prioritization output, stating an assumption that a detailed outcome of the prioritization step will be published for transparency. Another reviewer recommended that further testing and review of the tool be done within a public- and stakeholder-engaged process.

One reviewer suggested that the systematic post-market assessment process can be enhanced by adding an evaluation step following risk management action.

FDA Response:

We thank the reviewers for the additional suggestions/comments.

We reviewed the scoring definitions and did not add or remove score options for any of the Public Health criteria. However, we will indeed consider adding more score options (within 1-9) in future versions of the tool if the need arises. We did assess the potential impact on score differentiation if the current criteria scoring were changed to the standardized score matrices suggested by the peer reviewer. Using baseline weighting for overall score determination, if criteria scores are standardized to all be {1, 5, 9}, that would result in fewer unique overall scores, and if standardized to {1, 3, 5, 9}, there would be no change in the number of unique overall scores.

Discussion of transparency in the reviewer comments is included in response to Q1. We also added more information to the Overview section in the revised tool description document regarding transparency. As acknowledged in the revised tool description document, we anticipate that the prioritization tool will continue to be revised and updated as necessary (e.g., to remain current with evolving scientific knowledge, tools, and techniques). The results of the current peer review will be made publicly available alongside the updated and revised version of the prioritization tool. While we do not currently have plans to solicit peer review on future versions of the tool, we will consider this recommendation for future implementation.

It is not clear what the reviewer specifically intended by “an evaluation step after risk management,” but we agree that the systematic process will need to be routinely evaluated for its effectiveness as a program. This is a new program, and the need for specific revisions may arise once the process is applied and has been operationalized.