CENTER FOR Science IN THE Public Interest
Your Food and Health Watchdog

August 18, 2025

Martin Makary, MD, MPH
Commissioner of Food and Drugs
US Food and Drug Administration
10903 New Hampshire Avenue
Silver Spring, MD 20993

RE: FDA Tool for the Prioritization of Food Chemicals for Post-market Assessment (Docket No. FDA-2025-N-1733)

Dear Commissioner Makary,

The Center for Science in the Public Interest (CSPI) respectfully submits these comments in response to Docket No. FDA-2025-N-1733. We appreciate the agency's ongoing commitment to reforming the post-market assessment system for evaluating the safety of food chemicals, and we recognize the development of this tool represents important progress towards that goal.

In written comments we submitted to the agency in January 2025 regarding the discussion paper for an enhanced post-market assessment framework, we recommended that FDA develop and implement a risk-based system that would prioritize the riskiest chemicals.[1] The tool described in the current document should achieve that, as it approximates risk by incorporating scores for both hazard (toxicity) and exposure as well as by taking into consideration subpopulations that are particularly susceptible to toxic exposures. Consequently, we are generally in favor of FDA using the tool described for the prioritization of food chemicals for post-market assessment. However, we have several suggested revisions that will help ensure the tool produces prioritization rankings that are predominantly driven by risk and that such prioritization is sufficiently transparent and clearly described to promote public trust in the process.

Our top recommendations are:

- Provide greater weight to the Public Health Criteria score
- Adopt a continuous scoring system
- Revise toxicity criterion scoring
- Incorporate endocrine disruption into the toxicity rubric
- Define "New Scientific Information"
- Solicit input from the public on which chemicals to prioritize

See subsequent sections for additional information and responses to FDA's questions.

[1] Center for Science in the Public Interest. RE: Development of an Enhanced Systematic Process for the Food and Drug Administration's Post-Market Assessment of Chemicals in Food; Public Meeting; Request for Comments (Docket FDA-2024-N-3609). January 21, 2025. Available: https://www.cspi.org/resource/final-comment-re-development-enhanced-systematic-process-fdas-post-market-assessment.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

Top Recommendations

Provide Greater Weight to the Public Health Criteria Score: In order for this tool to be sufficiently risk-based, the Public Health Criteria should be the primary driver of the overall score, but the overly complicated system of scoring in the tool has created a situation where the Other Decisional Criteria can have an outsized influence on the overall score. This is the case because equal weights are applied to the two sub-scores (Public Health Criteria Score and Other Decisional Criteria Score) and because there are only three Other Decisional Criteria and four Public Health Criteria, meaning the individual contribution of one Other Decisional criterion to the overall score is weighted more heavily than the individual contribution of any one Public Health criterion. With the weighting applied as described, a 1-point score in a Public Health criterion contributes 0.125 points to the overall score, and 1-point score in an Other Decisional criterion contributes 0.167 points to the overall score. Thus, if Chemical A receives scores of 1 for all criteria except toxicity (a Public Health criterion), where it scores a 5, its overall score is 1.5. If Chemical B scores 1 for all criteria except stakeholder interest (an Other Decisional criterion), where it scores a 5, its overall score is 1.667. Thus, in this hypothetical scenario, the chemical with a lower toxicity score but higher score for stakeholder interest would be prioritized over a chemical with a higher toxicity score but a lower stakeholder interest score. This is not fully consistent with a risk-based approach to prioritization.

The simplest way to address this is to calculate the overall score by directly averaging all seven criteria (i.e., do not separately calculate the sub-scores for the Public Health Criteria and the Other Decisional Criteria to use in calculating the overall score). In effect, each criterion would receive a weight of 1/7, and because there are four Public Health Criteria and three Other Decisional Criteria, the Public Health Criteria would make up 4/7 of the overall score and the Other Decisional Criteria would make up 3/7 of the overall score. Thus, weighting all individual criteria equally and skipping the intermediate weighting step will ensure that the Public Health Criteria are not inappropriately outweighed by the Other Decisional Criteria. It is worth noting that under this revised approach to scoring, using the same hypothetical example as above, Chemical A and Chemical B would receive identical overall scores of 1.57. The agency should consider whether this outcome is in keeping with its goals of using a risk-based approach to prioritize chemicals, or if it is preferable to further revise the weighting system such that Public Health Criteria receive an even greater weight to ensure that Chemical A would receive a higher overall score than Chemical B.

Adopt a Continuous Scoring System: FDA intends to assign scores of 1, 3 (only for some criteria), 5, and 9 to each criterion. This approach allows a high score in one criterion to dwarf the others in calculating the overall score, but the equal weighting approach the agency has taken seems intended to put the individual criteria on more or less equal footing. The US Environmental Protection Agency's approach, upon which this tool is based, uses a different scheme, where the possible criteria scores are 1, 2, and 3.[2] We recommend that the agency

[2] US Environmental Protection Agency. TSCA Work Plan Chemicals: Methods Document. February 2012. Available: https://www.epa.gov/sites/default/files/2014-03/documents/work_plan_methods_document_web_final.pdf.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

instead use a scoring system where criteria scores are continuous (e.g., 1, 2, 3, and 4) and the top score is not inordinately high to avoid very high criteria scores, thus reducing the influence of any given criterion on the overall score.

Revise Toxicity Criterion Scoring: FDA states that a substance's overall toxicity criterion score will be the highest score in any of the seven stated categories of toxicological endpoints. On one hand, using this approach ensures that strong signals of severe toxicity result in the highest possible score, which is desirable, but on the other hand, this approach allows for little differentiation among chemicals with moderate or equivocal signals. Differentiation is potentially important in the mid-to-low range of toxicity criterion scores (i.e., it may be preferable that a chemical moderately associated with carcinogenicity and neurotoxicity receives a higher score than a chemical that is moderately associated with carcinogenicity but not neurotoxicity). Taking an average of the scores for each endpoint in the toxicity rubric might be one approach that could allow better differentiation, although we recognize the drawback that this could dilute the effect of strong toxicity evidence. Thus, we recommend that the agency consider whether the approach described will achieve the agency's goal of prioritizing chemicals that potentially pose the greatest risk and, if not, revise the toxicity scoring accordingly.

Incorporate Endocrine Disruption into the Toxicity Rubric: There is no mention of endocrine disruption in the toxicity rubric. Although the categories of toxicological endpoints already included could capture some adverse endpoints mediated by endocrine disruption (e.g., some reproductive, development, immunological, and neurotoxicological endpoints), endocrine disruption itself should be directly considered in the rubric. Perhaps this can be accomplished by adding a point(s) to the overall score when endocrine disruption is present. When considering endocrine disruption, the agency should not use dose thresholds to define scores—similar to what it proposes to do for immunotoxicity—because endocrine disrupting chemicals can elicit low-dose effects and demonstrate non-monotonic dose responses.[3,4]

Define "New Scientific Information:" The FDA should clarify the timeframes it will use to define what constitutes "new scientific information." The document states that "new scientific information" would be any information published since the last FDA evaluation of the substance. However, many substances have never undergone FDA evaluation because they came to market under the Generally Recognized as Safe (GRAS) loophole. We recommend the agency consider all information for such substances and that the agency compel industry to submit all relevant information purportedly substantiating GRAS status to the FDA, requesting additional authority if necessary.

[3] Melnick R, Lucier G, Wolfe M, Hall R, Stancel G, Prins G, Gallo M, Reuhl K, Ho SM, Brown T, Moore J, Leakey J, Haseman J, Kohn M. Summary of the National Toxicology Program's report of the endocrine disruptors low-dose peer review. Environ Health Perspect. 2002 Apr;110(4):427-31. https://doi.org/10.1289/ehp.02110427.
[4] Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR Jr, Lee DH, Shioda T, Soto AM, vom Saal FS, Welshons WV, Zoeller RT, Myers JP. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr Rev. 2012 Jun;33(3):378-455. https://doi.org/10.1210/er.2011-1050.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

Further, some food chemicals may have undergone FDA evaluations that were limited in scope or that were never made public. If the agency intends to use any such evaluation as the starting point for "new scientific information," that prior evaluation should be made public and there should be an opportunity for the public to provide input on it.

Solicit Input from the Public on Which Chemicals to Prioritize: FDA states that it will use various surveillance and signal detection tools to develop an inventory of chemicals to prioritize. CSPI is supportive of FDA using any available tools and technologies to identify relevant food chemical safety signals. In addition, we recommend that FDA solicit public input on which chemicals to include in that inventory, such as by issuing a request for information (RFI) or creating a nomination process, or the FDA could request public input when it releases the priority list of chemicals it eventually generates with this tool. We made a similar recommendation in our January 2025 comments. In those comments, we expressed concern that FDA's approach to signal detection would result in FDA only acting on signals that emerge in the future, whereas there are numerous existing safety signals that merit FDA attention now. This is important given that FDA's list of chemicals under review does not necessarily capture all of those existing safety signals.[5] For example, synthetic food dyes are not on that list[6] despite the fact that FDA has never publicly responded to the peer-reviewed systematic review published by the California Office of Environmental Health Hazard Assessment in 2021. That review concluded that synthetic food dyes "can cause or exacerbate neurobehavioral problems in some children" and explicitly called into question the validity of FDA's existing acceptable daily intakes (ADIs) for those dyes, which should have triggered a rigorous, transparent evaluation by FDA.[7] Gathering recommendations or nominations from the public would be one mechanism by which FDA could capture existing signals alongside new signals. Further, since FDA intends to incorporate actions by other government agencies into the Other Decisional Criteria scoring, it should also ensure that actions by other agencies also inform the creation of the chemical inventory in the first place. FDA should also specify how it will keep abreast of actions by other agencies, if it will solicit input from those agencies, and generally how it will be communicating with those agencies throughout the prioritization process. Soliciting public input will also help ensure alignment of the FDA's prioritization process with the priorities of external stakeholders, increase transparency, and improve public confidence in the process.

Responses to FDA Questions

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

[5] US Food and Drug Administration. List of Select Chemicals in the Food Supply Under FDA Review. Updated: June 18, 2025. https://www.fda.gov/food/food-chemical-safety/list-select-chemicals-food-supply-under-fda-review. Accessed: August 14, 2025.
[6] FDA's list does include FD&C Red No. 3, one of the synthetic dyes addressed by the OEHHA evaluation, but none of the others.
[7] California Office of Environmental Health Hazard Assessment. Health Effects Assessment: Potential Neurobehavioral Effects of Synthetic Food Dyes in Children. April 2021. https://oehha.ca.gov/media/downloads/risk-assessment/report/healthefftsassess041621.pdf.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

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. This overall approach seems appropriate and reasonable, with some potential modifications to clarify certain details. See recommendations in the preceding section and in response to other questions below for specific recommendations regarding weighting and scoring.

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. The Public Health Criteria overall seem appropriate in that they collectively qualitatively capture risk by incorporating toxicity (hazard), exposure, and considerations of impacts on particularly susceptible subpopulations. We particularly laud the agency for its inclusion of particularly susceptible subpopulations as a unique criterion.

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. The Other Decisional Criteria seem appropriate.

2.c. Are there additional criteria that should be considered? If so, please describe additional criteria that might be considered and why. FDA might consider revising the toxicity rubric and associated scoring to ensure that evidence of harm in humans results in a higher toxicity score than evidence of harm from other streams of evidence (animal studies, in vitro assays, etc.). This aligns with recommendations we made in our January 2025 comments.

3.a.i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why. The definitions for the Public Health Criteria are appropriate.

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. The toxicity criterion scoring is not entirely appropriate. We suggest that the toxicity criterion be revised to ensure that chemicals moderately associated with several toxic effects are prioritized alongside those that are strongly associated with a single toxic effect and that endocrine disrupting chemicals are appropriately captured by the toxicity score. See preceding section for additional details.

3.b.i. Are the definitions appropriately defined? If not, please describe changes that might be considered and why. The definitions for the Other Decisional Criteria are generally appropriate.

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? We recommend that FDA include additional details on how it will judge the risk of losing public confidence and assign scores of High, Moderate, and Low.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

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. Equal weighting of criteria within the Public Health Criteria seems appropriate.

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. Equal weighting within the Other Decisional Criteria seems appropriate.

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. As described in the preceding section, FDA's proposed approach to weighting and scoring creates a situation where the Other Decisional Criteria can have an outsized influence on the overall score. Above, we recommended that FDA calculate the overall score by directly averaging all seven criteria, such that all criteria receive a weight of 1/7, rather than calculating intermediate Public Health Criteria and Other Decisional Criteria scores. This should be the minimum revision the agency makes, but we argue that FDA could go further by assigning a higher weight to the Public Health criteria relative to the Other Decisional Criteria (though we did not make a specific recommendation about what those weights might be). FDA could develop and test several alternative weighting options using a small list of diverse chemicals. We would welcome an opportunity to review and comment on the results of such testing.

5.a. How might FDA incorporate information from new approach methodologies (NAMs) into the toxicity rubric? For chemical prioritization, we generally support the use of new approach methodologies (NAMs) as a means of hazard identification. As FDA is well aware, the validity and applicability of NAMs for human health risk assessment is an emerging area of research, and FDA and other US and international authorities are currently working to validate the use of NAMs in a regulatory context.[8] Thus, we recommend that decisions about the use of NAMs for chemical prioritization should be made in consultation with other US agencies and international authorities. The agency should consider using a Key Characteristics (KC) approach to evaluating and synthesizing data from NAMs and other mechanistic evidence. KCs are chemical and biological properties associated with certain classes of toxic chemicals. KCs have been established for carcinogens,[9] endocrine disrupting chemicals,[10] and immunotoxicants.[11] As part of its monographs program, the International Agency for Research on Cancer (IARC) uses the KCs of carcinogens to evaluate mechanistic data. The purpose of IARC's monographs program

[8] Interagency Coordinating Committee on the Validation of Alternative Methods. Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies. March 2024. https://doi.org/10.22427/NICEATM-2.
[9] International Agency for Research on Cancer. Preamble: IARC Monographs on the Identification of Carcinogenic Hazards to Humans. 2019. Available: https://monographs.iarc.who.int/iarc-monographs-preamble-preamble-to-the-iarc-monographs/.
[10] La Merrill MA, Vandenberg LN, Smith MT, et al. Consensus on the key characteristics of endocrine-disrupting chemicals as a basis for hazard identification. *Nat Rev Endocrinol*. 2020;16(1):45-57. https://doi.org/10.1038/s41574-019-0273-8.
[11] Germolec DR, Lebrec H, Anderson SE, et al. Consensus on the Key Characteristics of Immunotoxic Agents as a Basis for Hazard Identification. Environ Health Perspect. 2022;130(10):105001. https://doi.otg/10.1289/EHP10800.

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

is to identify carcinogenic hazards and evaluate the strength of hazard evidence, which is a somewhat analogous function to that of FDA's toxicity rubric. This comment should not be construed as CSPI endorsing the use of NAMs in other contexts, including the performance of human health risk assessments.

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. We do not recommend specific NAMs in this comment.

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? We suggest FDA consult with other US authorities, global authorities, and academic researchers. Above, we suggested FDA could consider revising its toxicity criterion scoring to ensure that evidence of harm in humans receives a higher score than evidence from other evidence streams, which might be a way to hedge against NAMs having an undue influence on the overall score, particularly for NAMs that have been shown to be useful for hazard identification but may not be useful or suitable for hazard characterization.

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? We support the use of structure-based screening tools to generate possible toxicity signals, but, as with NAMs, empirical evidence of a hazard should be given greater weight in scoring than structurally predicted toxicity. Again, taking stream of evidence into consideration for scoring toxicity would be one way to address this. This comment should not be construed as CSPI endorsing the use of Thresholds of Toxicological Concern (TTC) in other contexts.

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

- **Define "Sufficient Data**:" In scoring toxicity, FDA frequently mentions that it will consider whether sufficient data are available to evaluate the various endpoints in humans or animals. The document would benefit greatly from an explicit description of how FDA will define and assess data sufficiency.
- **Provide Rationale for Dose Thresholds**: For several toxicological endpoints, particularly for animal data, the FDA intends to use dose thresholds to define High, Moderate, or Low scores. We ask that the agency provide a rationale for these doses.
- **Define "Susceptible Subpopulation**:" We agree with FDA's inclusion of a criterion that accounts for susceptible subpopulations, and we ask that FDA specify how it will identify those subpopulations for each chemical. At a minimum, these subpopulations should include infants, children, and people who are pregnant or breastfeeding.
- **Revise Scoring for "External Stakeholder Attention:"** FDA indicates that a score of 3 ("uncertain due to conflicting attention") will be assigned on a scale extending to 9 when some stakeholders are giving the chemical attention while other stakeholders are not. It

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org

seems to us that this will be the case for the majority of chemicals of concern to others outside industry. Industry very often considers controversial chemicals to be safe, which might cause such chemicals to receive a score no higher than 3. We suggest the FDA revise its scoring to allow greater differentiation of scores in this criterion. Also, we recommend academia be listed among the stakeholders considered in this criterion. Finally, it seems possible that "External Stakeholder Attention" could be closely related to "Public Confidence" because, for example, attention to a specific substance from various stakeholders could drive public attention and concern or vice versa. We ask that FDA ensure that these two criteria are sufficiently differentiated such that interest from the general public is not scored twice.

- **Clarify Consideration of Non-Governmental Authorities**: There are several non-governmental authorities that are critical stakeholders in food chemical safety, including the FAO/WHO Joint Expert Committee on Food Additives (JECFA), the Codex Alimentarius Commission, and IARC. However, FDA's "Other Government Decisions" criterion only explicitly includes governmental agencies. We recommend that the scope of this criterion be expanded to include certain non-governmental authorities.

Concluding Remarks

CSPI is grateful that FDA is making progress towards reforming the federal food chemical regulatory systems and structures. We see evidence in this latest document that the agency considered comments CSPI and other stakeholders submitted earlier this year, and we are thankful for that. We look forward to additional opportunities to engage with FDA in its ongoing process to enhance the post-market framework for assessing food chemical safety.

Sincerely,

Thomas M. Galligan, PhD
Principal Scientist for Food Additives and Supplements
Center for Science in the Public Interest

1250 I Street, NW, Suite 500 • Washington DC 20005-5979 • tel 202 332-9110 • cspinet.org