

Response to Comments on the DfE Alternatives Assessment Criteria for Hazard Evaluation

August 8, 2011

The US EPA Design for the Environment (DfE) Program develops Alternatives Assessments, which are multi-stakeholder partnerships convened to evaluate priority chemicals and functional alternatives. Alternatives Assessments characterize chemical hazards based on a full range of human health and environmental information from the literature and models, and bring together stakeholders from across the spectrum of interested parties. Stakeholders help to shape the project scope, identify functional alternatives, help EPA consider economic realities, participate in development of data and the report, and are critical to implementation of safer alternatives. The outcome of these partnerships provides industry and other stakeholders with the information they need to choose safer chemicals, as well as avoid potential unintended consequences from switching to a poorly understood alternative.

DfE Alternatives Assessments are applied in situations where chemical hazard is a key impact in the lifecycle of a product or process. Alternatives Assessment is one of many tools that may be used to improve the sustainability of a product or process, and considerations such as cost and performance can complement this tool. The DfE Alternatives Assessment process is discussed on the DfE website at http://www.epa.gov/dfe/alternative_assessments.html.

Request for Comment on DfE Criteria for Alternatives Assessments

On November 30, 2010, the DfE Program posted on its web site for public comment the DfE Alternatives Assessment Criteria for Hazard Evaluation. The Alternatives Assessment Criteria use chemical hazard classification systems developed in the U.S. and internationally. For each toxicological endpoint, the criteria use the system that best differentiates between chemicals based on their hazard properties. To differentiate among chemicals based on hazard properties, the DfE Alternatives Assessment Criteria include endpoints that are likely to be satisfied with measured data or that can be modeled or addressed using established techniques supported by mature science. Structural analyses, models, and expert judgment – the tools of EPA’s New Chemicals Program – are used to satisfy endpoints in DfE’s criteria where data are not available. Additional endpoints may be satisfied for some chemicals, but alternatives can lack data and therefore, for endpoints where modeling is not possible, a meaningful comparison is difficult to make. If an endpoint is considered to be critical to understanding the chemical of potential concern, DfE will take measures to address it in an appropriate manner informed by input from stakeholders.

The Alternatives Assessment Criteria on which comment was sought and which this document addresses do not include the procedures for implementation of an Alternatives Assessment beyond assigning hazard values to toxicological endpoints.

DfE received 21 sets of comments, representing 55 named organizations or individuals (not including discrete trade association members), during the comment period, which ran from November 30, 2010 to January 31, 2011. Comments submitted on the criteria, the input of stakeholders involved in ongoing DfE alternatives assessments, and lessons learned from application of the draft criteria are incorporated

into the updated version of the Alternatives Assessment Criteria that accompanies this Response to Comments document. DfE thanks everyone who took time to submit comments, as well as those who have read the criteria and shared their support and ideas less formally.

Below, DfE presents and discusses the comments received on each provision of the criteria and indicates any planned changes to the proposed text. Please note that the comments have been paraphrased, summarized, and combined, as appropriate, for readability and to conserve space; full versions of the comments can be requested at DfE@epa.gov.

1. General Comments

Determining Which Chemical is Most Preferable

Comment: The Criteria do not describe how the endpoints would be weighted to determine which chemical(s) are most preferable. More guidance on interpretation and ranking is requested.

Comment: EPA should modify the assessment approach to require a summary that distinguishes better and worse alternatives based on toxicity and potential for human contact and/or environmental toxicity.

Response: DfE Alternatives Assessments are intended to provide the best information from testing and modeling to support decision making by the stakeholder community. In general, DfE does not plan to choose or impose a decision-making scheme. Organizations such as Clean Production Action have developed chemical scoring methodologies in partnership with industry sectors. Such methodologies may be used with the outcome of DfE Alternatives Assessments as a decision-making tool.

Comment: The summaries should address the impact of data gaps and scientific uncertainties to ensure that readers are guided to the most preferable alternatives.

Response: DfE Alternatives Assessments make explicit which hazard calls are based on data and which are based on models or expert judgment. This helps stakeholders recognize the uncertainties due to data gaps. Stakeholders and government agencies can incorporate this information into decision-making.

Updating DfE Alternatives Assessments

Comment: Will DfE update its AA reports with new or higher-tier data as they become available?

Response: DfE Alternatives Assessments cover a snapshot in time. EPA may choose to update a DfE Alternatives Assessment based on availability of new information or stakeholder interest, but the program does not anticipate periodic re-evaluations of its Alternatives Assessments.

Use of the European Union (EU) Dangerous Substances Directive (DSD) and the Regulation on Classification, Labeling, and Packaging (CLP)

Comment: DfE should not adopt the EU CLP or the DSD in its Alternatives Assessments.

Comment: The CLP has replaced the DSD. Therefore, the criteria should refer to the CLP not the DSD.

Comment: The Alternatives Assessment Criteria should expand the use of risk phrases and other classifications from the EU.

Response: Experience in implementing Alternatives Assessments has shown that EU classifications, along with classification from other sources (e.g., IARC, NIOSH), can enhance the evaluation of chemicals. The thresholds for the DSD's categories of danger, which were already included in the draft Alternatives Assessment Criteria, and the thresholds for CLP's hazard classes are generally similar to those in the DfE Alternatives Assessment Criteria. Therefore, DfE has expanded the criteria to include hazard statements from the CLP in addition to risk phrases from the DSD.

The criteria reference the CLP's hazard statements and the DSD's risk phrases because, under the CLP, a transitional period allows time for chemical suppliers to move from the DSD classification system to the CLP rules. Suppliers have until June 1, 2015 to complete the transition.

Relationship of DfE Alternatives Assessment to the Green Screen for Safer Chemicals

Comment: We agree with DfE's decision to harmonize with existing hazard evaluation protocols and regulations including the Globally Harmonized System for the Classification and Labeling of Chemicals (GHS). In addition, DfE should explicitly acknowledge relevant and similar chemical hazard assessment protocols, most notably Clean Production Action's Green Screen.

Response: The revised introduction to the Alternatives Assessment Criteria briefly describes how stakeholders might use the results of a DfE Alternatives Assessment. The Green Screen for Safer Chemicals is mentioned as one supplemental decision-making tool that could be used in conjunction with DfE Alternatives Assessments. The EPA DfE website also provides more information on Alternatives Assessment and related tools. Visit www.epa.gov/dfe/alternative_assessments.html for information and for a link to Clean Production Action's Green Screen.

Data gaps

Comment: A lack of data for an endpoint should be considered no more favorably than the worst hazard rating available for that endpoint. No data should be regarded as equivalent to the highest level of concern (e.g., High or Very High).

Response: DfE uses the tools, models, and scientists of EPA’s New Chemicals Program to apply analogs, models and expert judgment that allow for estimations of toxicity, persistence and bioaccumulation. The absence of test data cannot be assumed to be an indication of no concern and modeled data can be an important tool for prioritizing future testing and for ensuring that chemicals are differentiated based on the best data. In DfE Alternatives Assessment reports, including summary data tables, the Agency explains the source of information (test data or modeling). Use of expert judgment, models and analogs results in an estimated value and is generally associated with a lower level of confidence than determinations made based on measured data.

Expand the Toxicological Endpoints Considered

Comment: The criteria omit several important endpoints such as hepatotoxicity, nephrotoxicity, and hematotoxicity.

Response: The evaluation of repeated dose toxicity in the Alternatives Assessment Criteria is intended to address a wide range of possible impacts, including, but not limited to hepatotoxicity, nephrotoxicity, and hematotoxicity. Section 4.1.6 of the Alternatives Assessment Criteria addresses repeated dose toxicity. The testing required to address this endpoint, such as OECD Test Guideline 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents, includes routine clinical observations, hematology and clinical biochemistry, body weight/food and water consumption, as well as both gross necropsy and histopathology involving the following organs and organ systems:

Liver	Testes	Ovaries
Kidneys	Epididymides	Thymus
Adrenal glands	Uterus	Spleen
Brain	Heart	

In addition, tissue preservation for histopathology should include:

all gross lesions	skin and eyes	heart, aorta, trachea and lungs
brain (representative regions including cerebrum, cerebellum and	spinal cord (at three levels: cervical, mid-thoracic and lumbar)	bone marrow (and/or a fresh bone marrow aspirate)

medulla/pons)		
pituitary	thyroid and spleen	parathyroid
thymus	adrenals	gonads
uterus, prostate	accessory sex organs	female mammary gland
kidneys	urinary	bladder
liver	pancreas	small and large intestines (including Peyer's patches)
esophagus	salivary glands	stomach

The test methods further specify that clinical and other findings may suggest the need to examine additional tissues. Any organs considered likely to be target organs based on the known properties of the test substance should be preserved. Repeated dose testing is designed to be comprehensive in nature.

Comment: The environmental hazard traits should also be more comprehensive to include domestic animal toxicity; impacts on wildlife growth, survival, development and reproductive toxicity; loss of genetic diversity/biodiversity; and mobility in environmental media.

Response: The environmental criteria are focused on aquatic toxicity because these test data are most likely to be available for comparison, and robust models have been developed. Other data will be considered if it can be used to distinguish among alternatives. As placeholders for the eventuality that data and criteria are available, DfE has included domestic animal toxicity; impacts on wildlife growth, survival, development and reproductive toxicity; loss of genetic toxicity/biodiversity; and mobility in environmental media in the Additional Endpoints section of the Alternatives Assessment Criteria.

Comment: DfE should consider incorporating many of the same hazard traits and endpoints which have been designated by California's EPA, Office of Environmental Health Hazard Assessment (OEHHA) in their draft "Identification of Hazard Traits, Endpoints, and Other Relevant Data for Inclusion in the Toxics Information Clearinghouse in Green Chemistry Initiative."

Comment: DfE should add physical hazards to the AA criteria, both in the summary table and the narrative.

Response: A new section (#5), Additional Endpoints, has been added to the criteria document. This section includes a number of endpoints that could be added to an alternatives assessment if the majority of chemicals can be evaluated for that endpoint.

Add "Non-toxic" or "No concern" as a Possible Designation

Comment: We suggest that DfE add a "non-toxic" or "no concern" category to its Alternatives Assessment Criteria. These designations would be supported by evidence demonstrating that

the chemical is not toxic above a generally accepted limit dose, or that a chemical has been tested and demonstrated negative results.

Response: A designation of “non toxic” or “no concern” may be misleading to some audiences. Low hazard is an appropriate designation for endpoints for which chemicals pose a low concern for hazard.

Expand the Scope Beyond Drop-In Replacements

Comment: DfE should examine production alternatives instead of limiting scope to drop-in technologies. The Alternatives Assessment document can be modified to require an upfront assessment of the performance goal of the product in question (i.e. fire safe foam furniture or achievable sales records) and to identify any alternative processes that could achieve these goals while reducing the volume or toxicity of chemical additives, or reducing the potential for human exposure.

Response: The primary purpose of DfE Alternatives Assessments is to evaluate chemical alternatives. In the process of implementing most Alternatives Assessments, the stakeholder partnership considers alternatives beyond those that are chemical substitutions, and includes a discussion in the report. For example, a report on alternative developers for BPA in thermal paper might discuss alternative printing technologies and paperless (electronic) receipts.

Absorption, Distribution, Metabolism and Excretion (ADME)

Comment: ADME data should be incorporated into the Alternatives Assessment Criteria. Differences in ADME among species could be the reason why a chemical produces the same type of toxicity in multiple species, but at vastly different dose levels. ADME should be incorporated to ensure that classification is based on the species that is most “human” in their ADME response.

Response: DfE considers the impact of ADME data when they are available. In particular, DfE evaluates ADME data on metabolites of interest, such as those that are known to be more toxic, or that have structural alerts signifying concern. This possible use of ADME data is noted in section 2.6 of the Alternatives Assessment Criteria.

Degradation and Metabolism By-Products of Concern

Comment: How are degradation (e.g., biodegradation, thermal decomposition) and metabolism by-products considered?

Response: The degradation or metabolism of a chemical into a by-product which itself is hazardous, slow to degrade, or bioaccumulative will be included in the hazard assessment,

where such information is available. The intent is to capture degradation products or metabolites which contribute to the overall hazard potential of a chemical.

Emerging Science

Comment: How will ToxCast data be used in the hazard assessment?

Response: DfE, along with our colleagues within OPPT, will be engaged in extending the proof of concept of ToxCast data, by comparing ToxCast results with DfE hazard characterizations.

Comment: U.S. EPA should use an approach which defines early indicators of harm by considering early changes or perturbations as an adverse effect. This will be more consistent with recommendations of the NAS and will have the potential to decrease the amount of animal testing, while allowing for a better assessment of multiple chemicals and sensitive life stages.

Comment: How will epigenetic changes be addressed?

Response: Use of early indicators for hazard characterization, including epigenetic changes is an important goal, and DfE reviews and uses all available information when assessing chemicals. Evaluating perturbations is an emerging area of science, and is associated with some uncertainty.

2. General Requirements

2.1 Data for all relevant routes of exposure will be evaluated.

Comment: The AA Criteria state that data from all relevant routes of exposure will be evaluated. Are data from non-relevant exposure routes eliminated in the classification process?

Response: Well-performed studies using the most relevant routes of exposure are given the greatest weight of evidence in our assessments. Where additional data using atypical exposure routes are available, DfE will consider these data and their potential to add to the weight of evidence.

Comment: Only oral, dermal, and inhalation routes of exposure are mentioned in the criteria document. Additional relevant routes of exposure include transplacental transport, lactational transfer, as well as experimental animal models which use injection either as intraperitoneal or subcutaneous administration. All of these routes of exposure are relevant for fetal and neonatal exposure to chemicals during critical periods of development.

Response: DfE will include a statement that additional routes of exposure will be considered on a case-by-case basis. DfE recognizes that other routes of exposure are possible, including

transplacental transport, lactational transfer, and intraperitoneal or subcutaneous injection. Because these exposure routes are less common, DfE or other authoritative organizations have not established hazard thresholds and as a result, such data will be used to inform weight-of-evidence decisions.

2.2 Application of Effect Levels; and Potential Impacts to Vulnerable Populations and Life Stages

Comment: NOAEL and LOAEL values are entirely a function of study design. EPA should note that the criteria can reward studies that are intentionally tested at a higher dose. Therefore, EPA should not give the same weight to NOAELs and LOAELs in determining the hazard designation. GHS guidance speaks to the use of guidance values for a dose shown to produce a significant health effect, i.e., LOAELS.

Response: DfE interprets the relevance and meaning of NOAEL and LOAEL values on a case-by-case basis, similar to EPA's TSCA New Chemicals Program.

Comment: Will studies that do not follow Good Laboratory Practices (GLP) be considered? A ranking system to compare the strength of the evidence is suggested.

Response: Yes, any well-performed studies will be considered. Well-performed studies are those that use accepted experimental methods, relevant test systems, sufficient and applicable dosing protocols, and adequate dose groups and sizes, and include but are not limited to guideline studies. The hazard assessment will include an evaluation of the strength of the evidence used in making a hazard determination. Designations made on the basis of measured data will be denoted by bold, colored letters; designations made on the basis of estimated data will be denoted by black, italic letters.

Comment: EPA should provide further detail describing how "an assessment of potential impacts to vulnerable populations and life stages" will be conducted.

Response: The Agency's protective approach and emphasis on children's health require that significant weight be given to studies measuring potential impacts to vulnerable populations or life stages. Where data on children's health or vulnerable populations are available and are from well-performed studies, those data will influence hazard designations.

2.5 Sensitivity of Test Species or Strain

Comment: The document indicates that species/strains with known sensitivities should be given additional consideration when evaluating data. By inference, the reverse should also be considered, i.e., use of species/strains that are known to be less sensitive to a substance or group of substances.

Response: The Alternatives Assessment Criteria has been updated to reflect this point.

3. Terms

Add additional definitions

Comment: Please provide a definition of a “compound”.

Response: The updated criteria (Section 3, Terms) include a definition of a compound. This term is used interchangeably with “chemical”.

Comment: EPA should define what an “adverse effect” is.

Response: The EPA IRIS program has defined an “adverse effect”, and DfE has included this definition to the criteria document (see Section 3, Terms).

4. Toxicological Criteria

4.1.2 Carcinogenicity

Comment: Several commenters suggested that DfE reconsider the grouping of GHS Category 2, IARC Group 2B, EPA Group C, EU CMR Category 3 and EU Risk Phrase R40 (i.e., possible carcinogens) as substances with a high hazard for carcinogenicity. Commenters noted that under GHS, these substances are considered to have moderate hazard for carcinogenicity. These commenters also suggested that a designation of high concern should be reserved for risk to humans, not simply evidence of carcinogenicity in animals. This is the basis for high concern in GHS, and many other schemes.

Comment: Equivocal data should not be used to support a moderate ranking since by definition equivocal data are of uncertain significance.

Response: Based on these comments, DfE will modify the criteria. Positive results will now be broken into two categories, following the approach of IARC, GHS, and others. Chemicals that are known or presumed to be human carcinogens will be designated as Very High concern. Chemicals that are suspected to be carcinogenic to humans will be considered as High concern. DfE has eliminated the equivocal category; rather, moderate concern for carcinogenicity is indicated by inadequate or limited evidence of carcinogenicity in experimental animals (and inadequate in humans). This approach incorporates the suggestions received by the commenters and mirrors the International Agency for Research on Cancer’s (IARC) classification scheme.

Comment: The definition of carcinogenicity should include the development of malignant, benign, and pre-neoplastic lesions and should also include considerations for non-genotoxic and

non-mutagenic mechanisms of carcinogenesis such as hyperplasia, decreased apoptosis or epigenetic changes in gene expression. The International Agency for Research on Cancer (IARC) deems an agent carcinogenic if it is “capable of increasing the incidence of malignant neoplasms, reducing their latency, or increasing their severity or multiplicity.” The IARC preamble should be used to update the carcinogenicity criteria.

Response: DfE has updated the definition of a carcinogenic substance using the IARC preamble.

4.1.3 Mutagenicity/Genotoxicity

Comment: Several commenters suggested that DfE reconsider the grouping of GHS Category 2 and EU CMR Category 3 substances as high hazard for mutagenicity. It was noted that under GHS, these substances are considered to be moderate hazard for mutagenicity.

Response: Based on these comments, DfE will modify the criteria. First, the updated criteria add a distinction between germ cell mutagenicity and somatic cell mutagenicity. For germ cell mutagenicity, the criteria follow GHS criteria with positive results broken into two categories. Chemicals that are known to or regarded as if they induce heritable mutations in the germ cells of humans, e.g., GHS Categories 1A and 1B, will be considered Very High concern. Chemicals that can be classified as possibly inducing mutagenic effects in the germ cells of humans, GHS Category 2, will still be considered High concern for mutagenicity. Chemicals that show evidence of mutagenicity in somatic cells supported by both *in vitro* and *in vivo* studies will also be categorized as High concern. DfE has eliminated the equivocal category; instead, chemicals that show evidence of mutagenicity in somatic cells supported by positive results in *in vitro* or *in vivo* studies will be assigned a Moderate level of concern for mutagenicity.

4.1.4 Reproductive and Developmental Toxicity

Comment: Authoritative sources such as Cal Prop 65, the EU Risk Phrases and CMR List, and NIEHS CERHR reports should be included in the AA criteria. DfE should explain how designations by these sources would be used to classify chemicals under the AA criteria.

Response: Use of authoritative lists can expedite the evaluation of chemicals in a hazard assessment, and in many cases, classifications under authoritative lists correlate with the AA criteria. A new section in the criteria document describes how frequently-considered authoritative lists could be used to contribute to hazard evaluation in the AA criteria. In alternatives assessments conducted by DfE, EPA scientists evaluate the basis of the authoritative classification to verify that it is relevant to the alternatives assessment criteria.

Comment: DfE should clarify the reasoning for the proposed reproductive and developmental toxicity criteria, or alternatively, harmonize with the GHS system for these endpoints.

Response: The criteria for reproductive and developmental toxicity rely on thresholds developed by EPA's Office of Pollution Prevention and Toxics. The introduction of potency through the use of these dose thresholds provides greater power to discriminate among chemical alternatives. The GHS classifies chemicals as reproductive or developmental toxicants on the basis of evidence demonstrating that an effect is known, presumed or suspected to be relevant to humans. The potency of a chemical to cause reproductive or developmental effects is not specifically addressed in the GHS, although the authors acknowledged that a "limit dose" above which adverse effects are of lower concern is scientifically relevant. REACH Annex IV provides threshold criteria for chemicals with an absence of "significant toxicological effects", including criteria for reproductive and developmental toxicity. DfE has added this reproductive and developmental toxicity threshold to the Alternatives Assessment criteria, creating a Very Low category.

4.1.5 Neurotoxicity

Comment: More guidance regarding the interpretation and ranking of data from neurotoxicity studies is needed.

Response: In the Test Methods and Data Interpretation section of the criteria document are two data interpretation references. In particular, one of these references, EPA's *Guidelines for Neurotoxicity Risk Assessment*, gives extensive detail about the identification of neurotoxic effects.

Comment: Give a more detailed example of how the neurotoxicity (or repeated dose) criteria will be modified for studies of duration shorter or longer than 90 days.

Response: For studies shorter or longer than 90 days, the criteria are prorated based on the duration of the study. For example, the 90-day threshold for High oral toxicity would be 10 mg/kg/bw/day. For a 28-day study, the threshold would be 30 mg/kg/bw/day. This is the approach suggested in the GHS.

4.1.6 Repeated Dose Toxicity

Comment: DfE should clarify the reasoning for the proposed AA criteria for repeated dose toxicity, or alternatively, harmonize them with the GHS system.

Response: The criteria for repeated dose toxicity are taken from the GHS system. DfE has updated the criteria to clarify that this is the case.

Comment: In the GHS, it is noted that when considering repeated dose toxicity, it is important to assess the severity of effects observed when determining whether a chemical is classifiable as a

chronic toxicant. For repeated dose toxicity, the Alternatives Assessment Criteria should refer to the GHS Specific Target Organ Toxicity – Repeated Exposure chapter.

Response: Severity is one of the considerations for interpreting data during the hazard assessment. Severity is addressed in depth in the Test Methods and Data Interpretation section of the criteria document where DfE references the GHS chapter: Specific Target Organ Toxicity – Repeated Exposure.

4.1.7 Sensitization

Comment: DfE has not provided guidance on the characterization of respiratory sensitization. In other contexts, DfE has already established criteria for this endpoint, for example, in the criteria for fragrances, and there are several useful lists from which to draw.

Response: DfE will add a new section with criteria for respiratory sensitization based on GHS criteria. High, Moderate, or Low characterization of hazard will be included where data are available. DfE will also add an explanation that a chemical's presence on a list of respiratory sensitizers will be used to characterize hazard; the level of concern will be dependent on an evaluation of the strength of the data used to list that chemical.

4.1.9 Endocrine Activity

Comment: Endocrine activity should be included in the hazard comparison table with High/Moderate/Low-type hazard characterizations.

Comment: Inclusion of endocrine activity in the hazard comparison table is of critical importance. Endocrine activity should be included in the table either using the current characterizations or an H/M/L- scheme similar to the other endpoints.

Response: Data for evaluation of endocrine activity for most chemicals will be limited. For example, of the 19 alternatives to BPA in thermal paper evaluated thus far by DfE, only eight chemicals have relevant data. While all eight have *in vitro* data, only four have *in vivo* data. This makes it difficult to compare results. Data are limited for most chemicals, robust tools are not available for modeling endocrine activity and, as a result, describing the level of concern for hazard would not be distinguishing among chemicals evaluated. Because of these limitations, the criteria indicate that DfE will summarize the data in terms of positive or negative results, and relative potency (compared, for example, to estrogen and BPA). The approach proposed in our Criteria will provide a summary of available data. It is not feasible to provide this information in a hazard comparison table.

Comment: DfE has created a new definition that lacks support from the scientific literature. DfE should make a clearer distinction between endocrine activity and endocrine disruption. A better explanation as to why a hazard call will not be made is needed. Endocrine activity and endocrine disruption are modes of action, not an adverse effect. Consensus on the definition of

endocrine activity and its implementation should be required before DfE includes it in the Alternatives Assessment Criteria.

Response: Currently available data for most chemicals will allow only a summary of study results and a finding that results are positive, negative or equivocal for endocrine activity. Alternatives assessment reports will contain this level of description under “endocrine activity.” The approach will allow evaluation of the results of *in vitro* assays (such as receptor binding assays) and a determination of whether a chemical could interact with the endocrine system.

In the future, it may be appropriate to designate level of hazard for an endocrine endpoint. Such an approach would depend on development of test methods and decision criteria.

Endocrine activity is often considered a mode of action, similar to mutagenicity and genotoxicity and therefore it is appropriate to include endocrine activity in the Alternatives Assessment Criteria.

Comment: There are significant unresolved issues: EDSP has not achieved consensus on the testing requirements and combination of results that lead to the characterization of endocrine disruptor; regulatory programs acknowledge that the science is not ready for use in their decision-making. Until there is broad consensus within EPA and elsewhere, DfE should not include this criterion.

Response: The DfE Alternatives Assessment Criteria was developed with an awareness of the EDSP. DfE reports will include data from well-conducted studies published in the peer-reviewed literature, including effects related to endocrine systems.

Comment: Use of non-standard test guidelines poses challenges for interpretation. Guidance from EDSP on interpretation is recommended. DfE should include data on aquatic organisms. DfE should discuss the level of evidence for potential endocrine activity. There is currently no category for “clearly endocrine active”.

Response: As is the case for all endpoints, DfE considers both guideline and non-guideline studies to inform understanding of an endpoint. Studies will be considered in a weight-of-evidence determination and EPA experts in endocrinology will be helpful in interpreting data.

The criteria’s list of preferred test methods has been updated to include those developed under the EDSP. Methods for non-mammalian species will also be included. Positive *in vitro* binding assays are evidence of potential endocrine activity. As the science advances, the criteria could be updated to include a category for “clear evidence of endocrine activity.”

4.2 Environmental Toxicity and Fate

4.2.1 Aquatic Toxicity

Comment: Regulatory and international frameworks such as REACH, OECD, and GHS employ NOEC (or Ecx) values for chronic aquatic toxicity, while the draft DfE Alternatives Assessment Criteria employ LOEC. It's hard to find other frameworks where LOEC is employed.

Response: The EPA New Chemicals Program uses LOEC values in its assessments of new chemicals. For some studies, it's possible that only a LOEC or only a NOEC may be available. Therefore, DfE will modify the criteria to allow for consideration of both values. This approach is consistent with DfE's evaluation of both no and low adverse effects levels for human health toxicity.

Comment: The test methods section (5.8) could be supplemented with chronic aquatic toxicity test methods, for example, those from EPA OCSPP or OECD.

Response: We agree with the suggestion and have amended the criteria document to include these test methods.

Comment: The aquatic toxicity criteria should be consistent with the GHS system.

Response: The acute aquatic toxicity criteria are aligned with the GHS system. The GHS takes a multi-pronged approach for chronic aquatic toxicity, incorporating the potential for biodegradability or low bioaccumulation potential into its hazard categories. DfE has separated these hazard endpoints in the Alternatives Assessment Criteria. The chronic toxicity criteria rely solely on chronic data. Persistence and bioaccumulation are evaluated separately in other parts of the document.

4.2.2 Environmental Persistence

Comment: How will environmental monitoring data be used to modify a persistence designation? If a chemical does not pass a ready test, what data will be used to determine the half-life of the chemical?

Response: In cases where a chemical does not pass a ready test, DfE will search for additional experimental data to determine the half-life of the chemical. Data from simulation tests may be available, for example. In the absence of measured data on the substance of interest, DfE will evaluate data for suitable analogs and use estimated values from models such as EPI (Estimation Program Interface) Suite or SPARC (Sparc Performs Automated Reasoning in Chemistry).

Comment: Degradation products of concern should be explicitly integrated into the hazard assessment. The term degradation products of concern should be defined and a protocol for identifying such products be provided.

Response: DfE will clarify the persistence criteria and the manner in which degradation products will be considered. Under section 4.2.2, Environmental Persistence, DfE has changed “degradation products of concern” to “persistent degradation products”. It is DfE’s intent to evaluate persistence based on complete (ultimate) degradation, thereby considering the potential for persistent degradation products. Degradation or metabolism by-products that are toxic to humans or other organisms will be considered in the appropriate toxicity endpoint. DfE will explain, in Section 2 – General Requirements, how metabolites and degradation products will be considered.

Comment: The 10-day window should not be incorporated into the criteria, or at a minimum, the criteria should specify that the 10-day window applies only to pure/discrete organic substances and not technical mixtures.

Response: DfE follows OECD guidance on the applicability of the 10-day window to structurally similar mixtures. OECD states that if “a test on the mixture is performed and it is anticipated that a sequential biodegradation of the individual structures is taking place, then the 10-day window should not be applied to interpret the results of the test”. DfE will include a description of this guidance in the persistence criteria.

Comment: DfE states that for persistence evaluations, in the absence of measured data, analogs and estimations may fill data gaps. The persistence criteria are suitable for estimations generated by EPI Suite, but do not address how other data, such octanol/water partition coefficients or organic carbon/water partition coefficients may be used to inform persistence.

Response: EPI Suite uses octanol/water partition coefficients (specifically, log KOW) to generate estimates of persistence. When octanol/water partition coefficients are available, the experimental value can be used in the model. EPA will consider other data such as organic carbon/water partition coefficients that support a weight-of-evidence evaluation.

Comment: Additional guidance and clarification is needed in the AA criteria to describe the requirements for environmental persistence. For example, the relevant modes of degradation, in addition to biodegradation, should be noted (e.g. hydrolysis, photolysis, oxidation/reduction).

Response: The criteria note other methods of degradation that may be relevant to the assessment.

4.2.3 Bioaccumulation

Comment: Explain how environmental monitoring data will be evaluated under the criteria. Can biomagnification and trophic magnification be factored into the evaluation? How would these change the hazard designation? How would *in vitro* fish metabolism data be factored in?

Response: DfE will incorporate environmental monitoring data and measurements of bioaccumulation potential, such as biomagnification and trophic magnification, in the weight of evidence where data are available. Where available, *in vitro* fish metabolism data could be used in the BCFBAF model.

Comment: Explain why modeled data are requested to support experimentally low BAFs. Why is a measured log Kow < 2 not sufficient to demonstrate that a substance is not bioaccumulative? Measured log Koas are not usually available, and there is no guideline method.

Response: EPA agrees with the comment that experimental BAF values <2 and low measured log Kows (also < 2) can be sufficient to designate a chemical as Low concern for bioaccumulation under the proposed criteria. The criteria are updated to reflect this point. EPA is interested in whether a chemical has the potential to bioaccumulate in air-respiring organisms, as well as aquatic organisms. For this reason, DfE will evaluate log Kows (relevant for the aquatic environment), and log Koas (relevant for air-respiring organisms). While standard methods for measuring Koa may be lacking, this value can be estimated from octanol/water and air/water partition coefficients, and the model KOAWIN in EPI Suite can also be used.

Comment: The criteria for Very High should be adjusted so that they are consistent with the EU, Stockholm Convention, and other schemes.

Response: EPA has considered this comment, and agrees with this recommendation. We have modified the Very High criteria (from 100,000 to 5,000) so that they are now consistent with these other approaches.

Stakeholders also expressed interest in better understanding the need for a threshold below 1,000 in the bioaccumulation criteria and the choice of the threshold of 100 for low concern. In developing the draft criteria, DfE's goal was to create a spectrum of bioaccumulation potential to differentiate among chemicals. The Arnot & Gobas 2006 data showed that a significant percentage of subject chemicals had BCF/BAF values below 1,000, making further differentiation desirable. After publication of the draft criteria, DfE discussed this issue with several technical experts within and outside of EPA who supported the use of a threshold at 100, an order of magnitude below the moderate range threshold of 1,000 as a useful means of differentiating among chemicals.

Comment: Request reconsideration of the use of Arnot & Gobas 2006. Guidance from the European Chemicals Agency is in conflict with the assumptions of the Arnot-Gobas BAF-QSAR model. ECHA uses data on molecular size and weight to determine a chemical's potential to bioaccumulate. The data on biomagnification of PBDEs conflicts with expectations from the Arnot-Gobas BAF-QSAR model, namely, that it was not hydrophobicity, but molecular size and/or weight that restricted biomagnifications. REACH guidance has identified a direct

relationship between log K_{ow} and BCF. If a substance does not bioconcentrate in mammals, then it is likely not to bioconcentrate in fish. Recommend using data from mammalian studies to inform potential to bioaccumulate. EPA generally accepts that substances of high molecular weight are not accumulating, perhaps a cut-off value for molecular weight and/or diameter could be included.

Response: There is no consensus in the scientific community around the usefulness of molecular size cut-offs for bioaccumulation. For example, Arnot et al 2009¹ evaluated molecular size cut-offs for bioaccumulation described by ECHA, and concluded that these cut-offs “are not supported by a critical review of the available data.” DfE will consider case-specific information and stakeholder comment as it assigns values for bioaccumulation. DfE will use all available, well-conducted studies when evaluating bioaccumulation potential, including data on mammalian species.

¹ Arnot, J. A., M. I. Arnot, et al. (2009). "Molecular size cutoff criteria for screening bioaccumulation potential: Fact or fiction?" Integrated Environmental Assessment and Management 6(2): 210-224.