Mixtures Similarity Tool (MiST)

User Guide

03/09/22

CONTENTS

1.0 MiST Overview 7
1.1 Analyzing Toxicity of Mixtures 7
1.2 Theoretical Framework for MiST 8

2.0 MiST Quick Start Guide 13
2.1 Minimum Data Requirements 13
2.2 Optional Inputs 13
2.3 Quick Start Guide 14

3.0 TUTORIAL: Step-by-Step Guide to Analyzing Mixtures 19
3.1 Step 1: Setting Up the Tool 19
3.2 Step 2: Entering Mixture Properties 19
3.2.1 Entering a New Mixture or Editing a Mixture 20
3.2.2 Entering BMD Type 22
3.2.3 Entering the ED Value 24
3.2.4 Entering the Mass Fractions 25
3.2.5 Locking the Mixture 25
3.2.6 Additional Tools in the “Data Repository” Tab 26
3.3 Step 3: Determine the Parameters and Mixtures for Analysis 26
3.3.1 Select Equal or Unequal Potencies 27
3.3.2 Enter the Critical Value 28
3.3.3 Select Data Rich or Data Poor 29
3.3.4 Enter the Candidate Mixture Name 29
3.3.5 Enter the Number of Reference Mixtures 30
3.3.6 Enter the Reference Mixture Names 30
3.3.7 Run the Analysis 31
3.3.8 Additional Tools in the Settings Tab 31
3.4 Step 4: Review the Results 32
3.4.1 Tool Run Settings 32
3.4.2 Results for the Candidate Mixture 33
3.4.3 Results for the Reference Mixture 34

4.0 MiST Documentation 35
4.1 Assumptions 35
4.2 Characterizing the Uncertainty in Reference BMD 36
4.3 The Monte Carlo Approach 37
4.4 The Data Rich Case 38
4.5 The Data Poor Case 39

5.0 MiST Validation 40

6.0 Troubleshooting 43

7.0 Glossary 44

8.0 References 47

TABLE OF TABLES

Table 1. MiST input data availability scenarios....12
Table 2. Setup in the Quick Start Guide Tab....14
Table 3. Step 1 – Enter Mixture Properties in the Data Repository Tab....14
Table 4. Step 2 – Determine the parameters and mixtures for analysis in the settings tab....16
Table 5. Step 3 – Reviewing Results on the Results Tab....17
Table 6. Summary of Internal MiST Validation....40

TABLE OF FIGURES

Figure 1. Simplified version of the MiST approach that compares one candidate mixture to two reference mixtures....10
Figure 2. Decision tree illustrating how data inputs relate to available MiST comparison scenarios....12
Figure 3. Data repository tab for entering mixture properties....20
Figure 4. New mixture and unlock buttons....20
Figure 5. Specifying the mixture to unlock in the dialogue box....21
Figure 6. Entering the mixture name....21
Figure 7. Selecting the BMD type....22
Figure 8. Entering the ED value in the ED field....24
Figure 9. Enter mass fraction values....25
Figure 10. “Lock Which Mixture #?” dialogue box....26
Figure 11. New mixture, delete, and reorder features....26
Figure 12. Selecting equal or unequal potencies....27
Figure 13. Entering the CV....28
Figure 14. Defining inputs as data rich or data poor....29
Figure 15. Entering candidate mixture names....29
Figure 16. Entering number of reference mixtures....30
Figure 17. Entering reference mixture names....30
Figure 18. Additional tools in the ‘Settings’ tab....31
Figure 19. Reviewing the results on the ‘Results’ tab....32
Figure 20. Run Settings entered on the ‘Settings’ tab populate on the ‘Results’ tab....32
Figure 21. Candidate mixture results for BMD, ED, nearest reference, and distance Dw (mean)....33
Figure 22. Reference mixture results for BMD, delta, Dw (mean), Dw (upper 95th), conclusion, and rank....34
Figure 23. Fill and erase buttons for interpolating or clearing CDF values....36

AUTHORS, CONTRIBUTORS, AND REVIEWERS

Federal Authors

Geniece M. Lehmann, Ph.D.
Laura M. Carlson, Ph.D.
Jeff Gift, Ph.D.

U.S. EPA, Office of Research and Development
Center for Public Health and Environmental Assessment (CPHEA)
Research Triangle Park, NC 27711

J Allen Davis, M.S.
Krista Christensen, Ph.D.

U.S. EPA, Office of Research and Development
Center for Public Health and Environmental Assessment (CPHEA)
Washington, DC 20004

Contract Authors

Cara Henning, Ph.D.
Graham Glen, Ph.D.
Joanne Trgovcich, Ph.D.
Tao Hong, Ph.D.

ICF
Durham, NC 27713

Contractor Support

Steven Black, M.P.H.
Rachel O'Neal
Courtney Lemeris, B.A.
Wren Tracy, M.H.S.
Jeremy Frye, M.S.L.S.
Rachel McGill, B.S.
Nicole Vetter, M.L.S.
Leah West, B.S.
Raquel Silva, Ph.D.

ICF
Durham, NC 27713

Reviewers

Federal Technical Reviewers
Glenn Rice, Sc.D.

U.S. EPA, Office of Research and Development
Center for Public Health and Environmental Assessment (CPHEA)
Cincinnati, OH 45268

Rachel Shaffer, Ph.D.
Jessica Soto Hernandez
Samantha Jones, Ph.D.

U.S. EPA, Office of Research and Development
Center for Public Health and Environmental Assessment (CPHEA)
Washington, DC 20004

Janice Lee, Ph.D.
Kris Thayer, Ph.D.

U.S. EPA, Office of Research and Development
Center for Public Health and Environmental Assessment (CPHEA)
Research Triangle Park, NC 27711

Cynthia Rider, Ph.D.

National Toxicology Program
National Institute of Environmental Health Sciences
General Toxicology and Cancer Group
Research Triangle Park, NC 27709

External Peer Reviewers
The Mixtures Similarity Tool (MiST) and this user guide were subject to peer review by independent external experts convened by Versar, Inc. under contract with U.S. EPA (EP-C-17-023 TO 68 HERH21F0356).

Stuart Harrad, Ph.D.

University of Birmingham,
Edgbaston Birmingham B15 2TT UK

Keri C. Hornbuckle, Ph.D.

4114 Seamans Center for the Engineering Arts and Sciences
University of Iowa, Iowa City, Iowa 52242

John C. Lipscomb, Ph.D., DABT, FATS

Center for Toxicology and Environmental Health (CTEH), LLC
5120 Northshore Drive
North Little Rock, AR 72118

ABBREVIATIONS

BMC	benchmark concentration
BMD	benchmark dose
BMDS	Benchmark Dose Software
CDF	cumulative distribution function
CPHEA	Center for Public Health and Environmental Assessment
CV	critical value
Dw	average weighted Euclidean distance between reference mixture and candidate mixture
EC	effective concentration
ED	effective dose
EPA	United States Environmental Protection Agency
ERC	extra risk concentration
ERD	extra risk dose
IRIS	Integrated Risk Information System
MiST	Mixtures Similarity Tool
PCBs	polychlorinated biphenyls
POD	point of departure
RfC	inhalation reference concentration
RfD	oral reference dose
SD	standard deviation
SE	standard error
SEM	standard error of the mean
VBA	Visual Basic for Applications

1.0 MIST OVERVIEW

1.1 Analyzing Toxicity of Mixtures

The Mixtures Similarity Tool (MiST) is a Microsoft Excel® based tool. MiST automates the process of evaluating the degree to which chemical mixtures are toxicologically similar, and whether they are ***"sufficiently similar"***[1] for risk assessment purposes. The purpose of MiST is to support risk assessment of chemical mixtures relevant to human health. Invariably, humans are exposed to complex mixtures of chemicals in the environment. Toxicological information for chemical mixtures is diverse and often incomplete. EPA guidance recommends utilizing whole mixture dose-response data for risk assessment, when possible, rather than data on individual mixture components. In cases where toxicological information for a mixture is incomplete, EPA recommends that data from a "sufficiently similar" mixture may be used as an alternative (U.S. EPA, 2000). MiST was developed as a tool for researchers and risk assessors to ascertain the degree to which ***"candidate mixtures"*** and ***"reference mixtures"*** are similar. For MiST, "candidate mixtures" are not well-studied and might lack dose-response information, while "reference mixtures" are better-studied and have estimates of dose/exposure levels at which adverse effects occur, along with variance information for these estimates. These comparisons can be used as a decision aid to help judge whether mixtures are sufficiently similar for the purposes of the investigator (e.g., whether similarity is sufficient to allow for use of a reference mixture's dose-response data in a risk assessment of a candidate mixture). The development of approaches for assessing sufficient similarity is an active area of research; various methods have been applied to diverse chemical sets, including disinfection byproducts (Bull et al., 2009; Feder et al., 2009a; Feder et al., 2009b; Rice et al., 2009) and petroleum substances (Murray et al., 2013). MiST is based on a method published by Marshall et al. (2013), which has since been used to evaluate mixtures of pyrethroids and botanical supplements (Catlin et al., 2018; Ryan et al., 2019).

Polychlorinated biphenyls (PCBs) are a class of synthetic compounds characterized by a biphenyl structure with chlorine substitutions. There are 209 possible PCB congeners that vary according to the numbers and positions of the chlorine substitutions on the biphenyl molecule. PCB congeners differ not only structurally but also in their toxicity in humans and animals. PCBs were commercially synthesized as mixtures of congeners, but the composition of the originally produced PCB mixtures can differ substantially from mixtures humans may be exposed to in the environment due to processes such as degradation and weathering (U.S. EPA, 2019). In addition, environmental PCBs can also be inadvertently generated through combustion activities or manufacturing (e.g., waste incineration, or pigment paint production) (Mao et al., 2020; Prithiviraj and Chakraborty, 2020; Vorkamp, 2015). Most toxicological studies in animals are conducted using commercial PCB mixtures; there are few toxicological data to

[1] Terms included in the Glossary (Section 5) are indicated in ***bold italicized*** text upon the first use in the User Guide.

represent environmental mixtures of concern for human health risk assessment. MiST was developed to identify whether PCB mixtures with well-characterized dose-response information (i.e., reference mixtures) are “sufficiently similar” to a different mixture that might not have dose-response information for the same health effect (i.e., candidate mixture). If one or more reference mixtures are identified as sufficiently similar, their dose-response information could be used to assess risk for the candidate mixture. Application of a reference mixture’s dose-response information to assess risk for a candidate mixture is especially useful when dose-response data are unavailable for the candidate mixture, as is the case for most PCB mixtures found in the environment. It is expected that, for many MiST analyses, well-studied commercial PCB mixtures with abundant dose-response data (e.g., Aroclors) will serve as reference mixtures for comparisons with candidate mixtures of PCB congeners measured in the environment. This approach can be used with dose-response data for any health effect of interest. Toxicological studies of commercial PCB mixtures have evaluated effects in the following health outcome categories: cardiovascular, dermal, developmental, endocrine, gastrointestinal, hematopoietic, hepatobiliary, immune system, metabolic, musculoskeletal, nervous system, ocular, reproductive, respiratory, and urinary system (U.S. EPA, 2019).

1.2 Theoretical Framework for MiST

The theoretical framework underlying MiST is described in detail in Section 4.0. MiST (coded in Visual Basic for Applications [VBA]) implements a modification of a published mixtures similarity approach (Marshall et al., 2013). The amended method considers one candidate mixture and one or multiple reference mixture(s). The method is intended to answer two questions:

- Is a given reference mixture "sufficiently similar" to the candidate mixture such that dose-response data for the reference mixture could be used in a risk assessment of the candidate mixture?
- If more than one reference mixture is "sufficiently similar" to the candidate mixture, which reference mixture is most similar to the candidate mixture?

To support these conclusions, the method establishes the "distance" between two mixtures based on:

- ***Mass fraction*** for each PCB congener in the mixture (required for all mixtures)
- Estimated dose (or concentration) associated with an adverse level of response, along with a measure of variance (e.g., a standard deviation (SD) or cumulative distribution function (CDF)) for the estimate (required for all reference mixtures). MiST can be used with any dose or concentration associated with an adverse level of response if it is accompanied by variance information. ***Benchmark concentrations (BMCs)*** or benchmark doses (BMDs)[2] as well as ***extra risk***

[2] The BMC/BMD is an estimate of the dose associated with a benchmark response (BMR) that is considered to represent an appropriate response level for risk assessment purposes (e.g., 10% extra risk for dichotomous endpoints). See the U.S. EPA BMD Technical Guidance (U.S. EPA, 2012) and the U.S. EPA BMD Software (BMDS) User Guide (U.S. EPA, 2020) for additional information on deriving BMDs.

concentrations or doses (ERCs or ERDs) from EPA's ***CatReg software***[3] are examples of values that could be used. For simplicity, the term "BMD" is used in MiST and in this user guide to refer to any such value compatible with the tool. Details on how to characterize BMD variance in MiST are provided later in this section.

- (Optional) The ***toxicological potency*** of each PCB congener relative to the potency of an index congener (often, the most toxic congener) (e.g., Holland and Pessah, 2021; U.S. EPA, 2010).

Following the example provided by Marshall et al., 2013, MiST calculates the average weighted Euclidean distance ***(Dw)*** between the BMDs for a reference mixture and the candidate mixture. Euclidean distance is a measure of the shortest distance between two points that incorporates, in this case, all 209 dimensions for the union of 209 potential constituents of a PCB mixture. In this method, a mixture's composition is represented by a vector in C dimensions, where C is the number of mixture components (so, for PCBs, 209 dimensions, see Figure 1). The Euclidean distance can be weighted to reduce or magnify the effect of some components of the mixtures based on information regarding potency. A criterion is established to determine which distances represent sufficient similarity, and which do not.

For each candidate-reference comparison, the first step is to compare the upper 95th percentile of all the estimates of D_w (D_wU_{95}) to an established distance bound, referred to as the ***"critical value" (CV).*** The CV defines the boundary of a similarity region for a given reference mixture's BMD. If the D_wU_{95} estimate for a candidate mixture falls within that boundary, the candidate mixture can be considered toxicologically similar to the reference mixture. If several reference mixtures are compared to the same candidate mixture, then the reference mixtures are ranked by their mean distance, and the reference mixture with the smallest mean distance is distinguished as the best match to the candidate mixture.

The CV should be determined before the analysis. Since the CV represents a similarity threshold, its determination should be based on statistical and biological dataset characteristics underpinning mixture similarity (e.g., variability of the observed responses, statistical and biological significance of differences in outcomes or outcome levels). The lower (BMDL) and upper (BMDU) bounds of the BMD confidence interval, as well as the BMD's CDF, if available, can help to inform the CV. Typically, when BMD estimates are available, the CV is calculated from the absolute value of the difference between the mixture's BMD and a user-defined mixture-specific ***effective dose (ED) (or effective concentration (EC))***.[4] Since the CV is meant to represent a statistically and biologically meaningful mixture-specific difference from the BMD, the appropriate ED may vary with each mixture (depending on, e.g., the estimated variance of the

[3] Please consult the CatReg User Guide for additional information on ERCs/ERDs (U.S. EPA, 2017).

[4] MiST can accommodate effect levels expressed as doses (e.g., BMD, ED, ERD) or concentrations (e.g., BMC, EC, ERC) as long as the units are consistent across all effect levels included within an analysis (i.e., a BMD should not be used with an EC and vice versa). For simplicity, the terms "BMD" and "ED" are used in MiST and in this user guide to refer to any type of effect levels compatible with the tool.

mixture's BMD)[5]. ED selection requires expert judgment (Marshall et al., 2013).

In MiST, the original method from Marshall et al. (2013) was modified to allow for Monte-Carlo estimation of the 95th percentile of the Dw. Section 5 provides additional details on the approach. Both the original and the modified approaches report a distance that is expected to be greater than the "true" distance in 95% of replicates if the uncertainty in the data were eliminated. Using the Monte-Carlo approach offers an advantage because it does not rely on the assumption that distance estimates are normally distributed. It allows for the use of the CDF produced by EPA's Benchmark Dose Software (BMDS) (U.S. EPA, 2012), which represents the BMD as the median or maximum likelihood estimate (MLE) of a profile likelihood distribution to describe the BMD variance. Each iteration produces one estimate of the distance (Dw) between the reference mixture and the candidate mixture BMDs. The tool performs 10,000 iterations for each reference-candidate comparison and selects the 95th percentile from this distribution (DwU_{95}). A graphical depiction of the method is presented in Figure 1Figure 1. Simplified version of the MiST approach that compares one candidate mixture to two reference mixtures

.

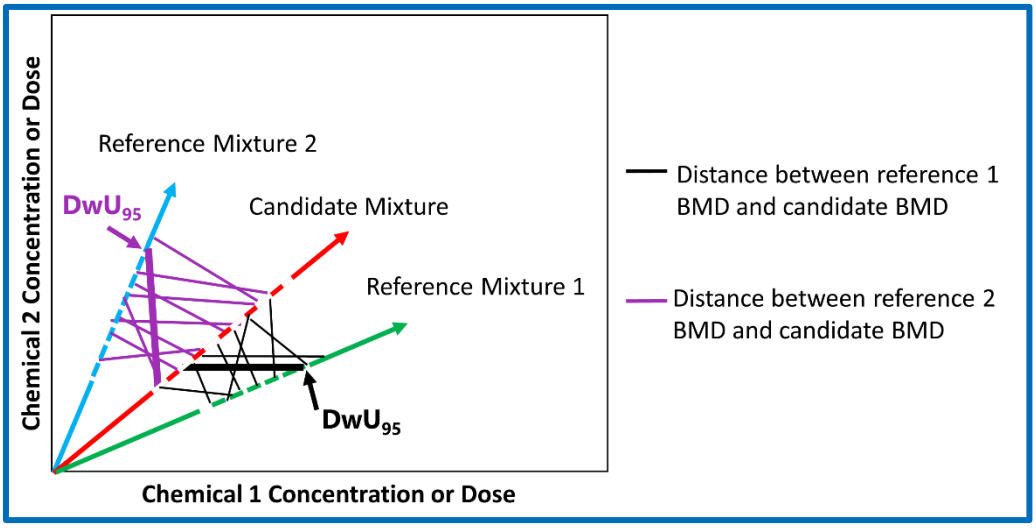

Figure 1. Simplified version of the MiST approach that compares one candidate mixture to two reference mixtures

The length of each dashed line segment reflects the uncertainty of the BMD for the candidate (red) and reference (blue and green) mixtures. The lines between the dashes represent thousands of Monte Carlo iterations performed to estimate distances (D_w) between the candidate BMD and the BMDs for reference 1 (black lines) and reference 2 (purple lines). Mixtures are considered similar if the upper 95th percentile for the Monte Carlo iterations (D_wU_{95} (bold lines)) is less than the critical value (CV)

[5] To maximize the tool's flexibility, instead of using separate CVs derived from a mixture's BMD and ED, MiST users also have the option to assign a single endpoint-specific CV that would be used for all candidate-reference comparisons. This approach does not consider the slope of mixture's dose-response curve and is generally not recommended for use at this time. However, it is available in case a future application is identified in which the selection of a single endpoint-specific CV is justified based on user objectives and data availability.

assigned to each reference versus candidate mixture comparison. This is a simplified two-dimensional depiction of the approach that has been extended in MiST to a 209-dimensional calculation for applications to PCB mixtures.

- MiST is designed to measure the similarity of a candidate mixture to one or more reference mixtures for different data availability scenarios. These scenarios, summarized in

Table 1, are based on the availability of information on the toxicological potencies of individual mixture constituents (e.g., PCB congeners) and dose-response information for the candidate mixture:

- For PCBs, relative ***toxicological potencies*** vary by health effect and are specific for each PCB congener. If congener-specific relative toxicological potencies are available for the health endpoint being considered (e.g., Holland and Pessah, 2021; U.S. EPA, 2010), they can be entered into MiST using the unequal potency option. For a given endpoint-specific analysis, the same congener potencies are used for all mixtures. If the toxicological potencies of the congeners are unknown, MiST assumes that all congener potencies are equal (i.e., all congeners have a relative toxicological potency estimate of one).
- In the ***data poor*** scenario (the principal scenario that MiST is intended to address), BMD information is not available for the candidate mixture but is available for the reference mixture(s). In this case, sufficient similarity is determined based on (1) the BMD distributions provided for the reference mixture(s) or estimated for the candidate mixture[6], (2) mixture composition as defined by the congener mass fractions, and (3) congener-specific potency information (if available). In *the* ***data rich*** scenario, both the candidate and the reference mixtures have dose-response information; MiST will estimate the overall distance between the candidate and reference mixtures using the dose-response information from each.

Note This User Guide focuses on the data poor scenario, as this is expected to be more commonly used and of most interest to users. In the more atypical data rich scenario and using the generally preferred option for CV selection, the default will be to make comparison(s) based on the difference between the BMD and ED for either the candidate mixture or reference mixture, whichever difference is larger. However, based on factors such as confidence in the modeling fit, dataset quality, and confidence in mixture composition characterization, users may wish to override this automated selection (e.g., to use the smaller difference) by specifying the CV to be used for the comparison in the Settings tab. Users should be prepared to justify the rationales for reference/candidate mixture designations based on their objectives and data availability.

- The dose-response information entered into MiST can take two different forms: either a BMD mean and SD or a dose-response CDF derived from EPA's BMDS (U.S. EPA, 2012). CDFs are included in the output of all model runs generated by the more recent versions of BMDS (i.e., starting with version 3.1.1). The different choices for the relative toxicological potencies, form of the BMD, and data

[6] Additional details for how the candidate mixture BMDs are estimated in the data poor scenario are described in Marshall et al. (2013).

scenarios define the overall six scenarios in the tool. A decision tree is provided in Figure 2

6	Equal	No BMD	Data Poor

[a] "Equal potency" refers to an unweighted analysis in which all congeners are assigned a weight of one; "unequal potencies" refers to a weighted analysis in which relative potency estimates for individual congeners are available and used. The same congener potency estimates are used for all mixtures in a given analysis (see section 3.3.1).

- to aid users in relating available input data to the appropriate MiST run scenario.

Table 1. MiST input data availability scenarios

Case	Toxicological Potencies[a]	Form of BMD for the Candidate Mixture	Data Scenario
1	Unequal	BMD as CDF	Data Rich
2	Unequal	BMD as mean ± SD	Data Rich
3	Unequal	No BMD	Data Poor
4	Equal	BMD as CDF	Data Rich
5	Equal	BMD as mean ± SD	Data Rich
6	Equal	No BMD	Data Poor

[a] "Equal potency" refers to an unweighted analysis in which all congeners are assigned a weight of one; "unequal potencies" refers to a weighted analysis in which relative potency estimates for individual congeners are available and used. The same congener potency estimates are used for all mixtures in a given analysis (see section 3.3.1).

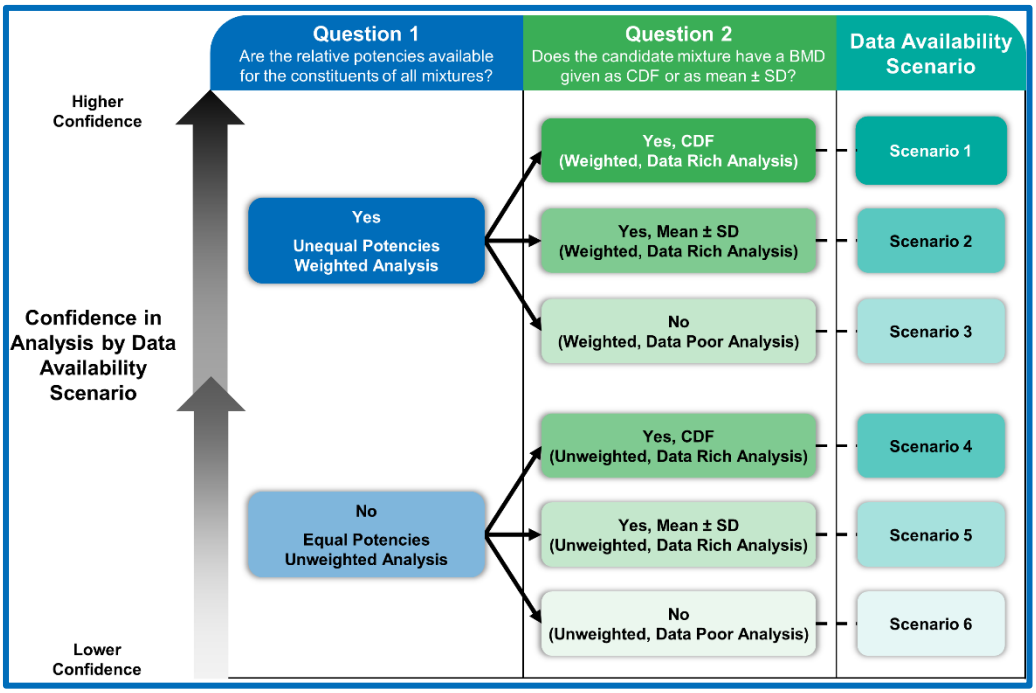

Figure 2. Decision tree illustrating how data inputs relate to available MiST comparison scenarios

The scenario numbers shown on the right side of the figure correspond to cases in

Table 1. Confidence in the analysis for these data scenarios is depicted by the color intensity of the boxes for each scenario and the gray arrows shown on the left side of the figure. Confidence is highest for weighted analyses in which the variance in the BMD for the candidate mixtures is available as shown at the top of the tree. Confidence in the analyses generally decreases with data availability scenarios lower in the tree.

2.0 MiST QUICK START GUIDE

2.1 Minimum Data Requirements

- Congener-specific mass fractions for all mixtures. The method by Marshall et al. (2013) uses Stork's methodology (Stork et al., 2008) for hypothesis testing for "equivalence" and mixed model theory to define sufficient similarity in dose-response for chemical mixtures containing constituents with different mixing ratios. MiST uses mass fraction to define a PCB congener's mixing ratio. A mass fraction is needed in MiST for each of the 209 congeners. However, users do not have to calculate mass fractions when mass or concentration data are available. If mass or concentration data are entered, the tool will adjust the values so that they sum to one (1), representing mass fractions. Congeners absent from a mixture have a mass fraction of zero. MiST was not designed to be used to compare partially characterized PCB mixtures, and conducting a MiST analysis with incomplete mass fraction data is not recommended. The mass fraction of any given congener can be calculated as follows:

$$w_i = \left(\frac{m_i}{m_T}\right)$$

where w_i is the mass fraction of the i^{th} congener;
m_i is the mass of the i^{th} congener;
and m_T is the mass of all congeners:

$$\sum_{i=1}^{209} = m_1 + m_2 + \cdots + m_{209}$$

- A BMD with uncertainty measures (i.e., SD or CDF) for at least one reference mixture.

2.2 Optional Inputs

- Critical value to compare to the 95th percentile of the distance.

Note The preferred approach is to allow MiST to calculate a CV for each mixture based on mixture-specific EDs and BMDs (when available), which is the default approach used in MiST if no CV is entered (see Section 1.2). The user has the option to assign a CV if needed but should not unless there is an appropriate justification for doing so.

- Relative toxicological potencies for mixture constituents. If these are unavailable, MiST assumes equal potencies.
- A BMD with uncertainly measures (i.e., SD or CDF) for all mixtures (i.e., "data rich scenario").

2.3 Quick Start Guide

A simplified guide for using MiST is presented in *Because the application of the methods described here and the interpretation of the results can be technically challenging, it is recommended that analyses be performed by or in collaboration with personnel expert in the procedures and potential pitfalls of this type of analysis*.

Note There is a tab within MiST named “Example Data”. These data are provided for new users who wish to gain experience with MiST data entry and functionality prior to entering their own data. It is intended to serve only as an example.

Table 2 below. Additional information for each step is provided in the step-by-step tutorial in Section 3.0. *Because the application of the methods described here and the interpretation of the results can be technically challenging, it is recommended that analyses be performed by or in collaboration with personnel expert in the procedures and potential pitfalls of this type of analysis.*

Note There is a tab within MiST named “Example Data”. These data are provided for new users who wish to gain experience with MiST data entry and functionality prior to entering their own data. It is intended to serve only as an example.

Table 2. Setup in the Quick Start Guide Tab

Quick Start Guide Tab	
Action	**Details**
Make sure that macros are enabled in Excel.	Either click the "Enable Content" button at the top when opening Excel or: 1. Click the File tab 2. Click Options 3. Click Trust Center 4. Click Trust Center Settings 5. Click Macro Settings 6. Select "Enable all macros”

Table 3. Step 1 – Enter Mixture Properties in the Data Repository Tab

Data Repository Tab	
Action	**Details**
1. Select a mixture to edit by clicking the "Unlock" button.	When prompted "Unlock Which Mix?", enter the mixture number. Cells will turn blue when unlocked. ***Note:*** When entering a series of new mixtures, users should select the button to add new mixture (s) for the intended analyses prior to entering data in the 'Settings Tab'. Once new mixtures are entered, users can select mixtures to edit/enter mixture properties.
2. Enter the name of the mixture.	Mixture # populates automatically ***Note:*** The name will be used in the 'Settings' tab to run the analysis.

Data Repository Tab	
Action	**Details**
3. Select one of three BMD types: a. CDF b. Mean & SD c. None	After selecting the BMD type, the tool will direct the user to input either the CDF or Mean & SD for the mixture, below the list of mass fractions. ***Note:*** When copying and pasting data, users should select the top cell where data will be pasted before copying and pasting data into that cell.
Table continues on the next page	
3a. Select one of three BMD types: **a. CDF** b. Mean & SD c. None	**CDF** If the BMD is entered as a CDF for this mixture, the user has two options: 1. In the 'Data Repository' tab, manually enter the BMD values in relation to their corresponding percentile. 2. In the 'Clipboard' tab, paste the percentiles and BMD values columns A and B, respectively (blue shaded columns). a. Return to the 'Data Repository' tab and click the "Copy CDF" button. b. When prompted 'CDF for which mix?', enter the mixture number. c. When prompted 'Is CDF data on clipboard?', select yes. d. The BMD values will then be pasted into the table below the mixture indicated above. ***Note:*** The BMD values should be entered as a whole number or decimal greater than zero in ascending order. ***Note:*** If a user does not have all of the BMD values for each percentile of the CDF for the mixture, a piecewise linear interpolation function can be performed to estimate the missing values. When manually imputing the BMD values for each percentile, enter as many values as possible and select the 'Fill' button located next to the lower CDF table. Using the 'Erase' button here will delete any values in the CDF table. This will automatically occur if the "Copy CDF" button is used. See section 4.2 for more information on using the Fill and Erase buttons to interpolate BMD values as a CDF. Additional information is available in section 4.2. ***Note:*** If not all CDF values are available for pasting in the clipboard, the user should add values for the percentiles that are available in consecutive lines. The "Copy CDF" and "fill" functions will assign values to the correct percentiles in the Data Repository tab. ***Note:*** Only blue cells can be edited. ***Note:*** Fill and Erase buttons are only available for unlocked mixtures.
3b. Select one of three BMD types: a. CDF **b. Mean & SD** c. None	**Mean & SD** Enter the mean and standard deviation to be used as the distribution of the mixture's BMD.
3c. Select one of three BMD types: a. CDF b. Mean & SD **c. None**	**None** Information on the BMD for this mixture will not be provided. ***Note:*** Mixtures with a BMD type of 'none' cannot be used as reference mixtures for the analysis.
4. Enter the ED value	***Note:*** The ED value should be entered as a whole number or decimal greater than zero. ***Note:*** Entering the ED is the preferred approach. Another option is to enter a CV in the 'Settings' tab. See sections 3.2.3 and 3.3.2 for more information.

Data Repository Tab	
Action	**Details**
5. Enter the masses, concentrations, or mass fractions of the congeners in each mixture	Enter the mass, concentration, or mass fraction of each congener in the mixture. If mass or concentration data are entered, the tool will adjust the values so that they sum to one (1), representing mass fractions. ***Note:*** Missing values will result in run errors. Thus, a value (which may be zero) must be entered for each congener. ***Note:*** If the user encounters a runtime error while entering data, click the "end" rather than the "debug" button. ***Note:*** MiST was not designed to be used to compare partially characterized PCB mixtures, and conducting a MiST analysis with incomplete mass fraction data is not recommended.
6. "Lock" the mixture.	When the information is inputted correctly for the mixture, select the "Lock" button at the top of the page and enter the mixture number in the pop-up box. This will lock the mixture and make it available for analysis.
Additional Tools in Repository Tab	**New Mixture** – Generates a new, blank mixture entry **Delete** – Allows the user to delete a mixture entry; based on the mixture number **Reorder** – Allows the user to reorder the mixture entries; the user is prompted to indicate the mixture to move by reporting the mixture number. The user will then be prompted to select the position the mixture should be moved to.

Table 4. Step 2 – Determine the parameters and mixtures for analysis in the settings tab

Settings Tab	
Action	**Details**
1. Select equal or unequal potencies	**Equal** – Signifies that the contributions of the mixture components are all treated equally. If using equal potencies, the relative toxicological potency estimate for each congener will be set to 1. ***Note:*** If an analysis is run with equal potencies, then the column of unequal potencies does not need to be filled. The sum of the relative toxicological potency estimates will be 1. **Unequal** – Signifies that relative toxicological potency estimates are entered for each congener (described in section 3.3.1). ***Note:*** If using unequal potencies, enter the relative potency estimates in the indicated column. If users select unequal potencies but do not enter potency data, then an error message appears. In accordance with the method described by Marshall et al. (2013), potencies are constrained to sum to the total number of components; for MiST, the total is 209 (corresponding to the total number of PCB congeners). This constraint makes the results of the equal and unequal potency analyses more directly comparable. If the supplied relative toxicological potency estimates do not sum to 209, the tool will autoscale the values so that the sum equals 209 once the “run analysis” action is initiated; autoscaled values will be displayed on the Results and Settings tabs (see section 3.3.1 for more information).
2. Enter the CV (optional)	The user specified CV should be entered as a whole number or decimal greater than zero. ***Note:*** Users must enter either ED OR a CV, but not both. Entering the ED is the preferred approach. See sections 3.2.3 and 3.3.2 for more information.

Settings Tab	
Action	**Details**
3. Select data rich or data poor	**Data rich** – Indicates that the BMD for the candidate mixture is available. **Data poor** – Indicates that the BMD for the candidate mixture is NOT available. Data rich analyses can be useful, for example, when a mixture has some dose-response information but is otherwise relatively poorly studied. It might then be compared with another, better-studied reference mixture to assess sufficient similarity, using not only the information considered in a data poor analysis but also the dose-response information that is available.
4. Enter the Candidate Mixture name	Determine which mixture from the 'Data Repository' tab will act as the candidate mixture for the analysis. Enter the name for this mixture in full. Mixture names are not case sensitive.
5. Enter the number of Reference Mixtures	Enter a positive, whole number of reference mixture(s) to compare to the candidate mixture. The appropriate number of boxes will pop up in column H. No more than 99 reference mixtures are allowed. ***Note:*** Reference mixtures must always have a BMD entered in the 'Data Repository' tab. BMDs are also required for candidate mixtures in a data rich comparison.
6. Enter Reference Mixture names	Determine which mixture from the 'Data Repository' tab will act as the candidate mixture for the analysis. Enter the name for this mixture in full. ***Note:*** The name is checked for validity only when another cell is selected. While mixture names are not case sensitive, the names must be spelled correctly in the Settings and Data Repository tabs.
7. Run the analysis	When the information is inputted correctly for the analysis, select the "Run Analysis" button at the top of the page. A successful analysis will result in a message that says, "Analysis Completed." ***Note:*** Microsoft excel disables the "undo" feature after running macros. Users are encouraged to save a version of the MiST workbook prior to running the analysis.
Additional Tools in this Tab	**Clear** – Clears the information from the 'Settings' tab; the 'Data Repository' tab will not be changed. **Save** – Allows the user to save the workbook; recommend saving each analysis as its own workbook.

Table 5. Step 3 – Reviewing Results on the Results Tab

Results Tab	
Action	**Details**
1. Run Settings	The settings selected for the analysis are returned here.

Results Tab	
Action	**Details**
2. Results for the Candidate Mixture	**Candidate Mixture** – The candidate mixture name is returned. **BMD^2** – The BMD estimate for the candidate mixture is returned. ED – The ED associated with a biological effect of the candidate mixture is returned (data rich runs only). **Nearest Reference** – Indicates the name of the reference mixture that is the most similar to the candidate mixture as determined by calculating and comparing the Euclidean distance between the candidate mixture and each reference mixture. **Distance *Dw* (mean)** – The average Euclidean distance between the nearest reference mixture (indicated above) and the candidate mixture**.** ***Note:*** If there are no reference mixtures deemed similar to the candidate mixture (similar reference mixtures have a 95th percentile Euclidean distance lower than the CV), the Nearest Reference will return 'None' and the Distance ***Dw*** (mean) will be left blank. The ***Dw*** for each Reference Mixture is returned in the results tab, regardless of the similarity conclusion (see item 3 below).
3. Results for the Reference Mixture(s)	**Reference Mixture** – The reference mixture name is returned. **BMD^2** – The BMD of the indicated reference mixture is returned. **ED –** The ED associated with a biological effect of the indicated reference mixture is returned. **Delta** – The CV for the analysis; set to |ED - BMD| by default[7]. Can also be set by the user with appropriate justification. ***Dw* (Mean)** – The average Euclidean distance between the reference mixture and the candidate mixture. ***Dw* (Upper 95th)** – The 95th percentile of the Euclidean distance between the reference mixture and the candidate mixture. **Conclusions:** ***Acceptable* –** *The **Dw** (Upper 95th) value for the reference mixture is less than the CV (Delta); consider the reference mixture as **SIMILAR** to the candidate mixture. The null hypothesis (that the mixtures have different toxicological effects) is rejected**.* ***Not Acceptable*** *– The **Dw** (Upper 95th) value for the reference mixture is greater than the CV (Delta); do **NOT** consider the reference mixture as sufficiently similar to the candidate mixture. The data do not support rejection of the null hypothesis.* **Rank:** when several reference mixtures are considered similar to the candidate mixture, the rank indicates which reference mixture is the best match to the candidate mixture. This is determined by comparing the ***Dw*** *(Mean)* for each reference mixture. The reference mixture with the smallest ***Dw*** *(Mean)* is considered the best match and assigned the rank of 1.
4. Use the clipboard to copy and store data for the Data Repository Tab	The clipboard allows users to store data for MiST runs. It also ensures all values are in an acceptable format for pasting into the Data Repository tab.

[7] For a data rich comparison, this can be based on either the reference mixture or candidate mixture, depending on which mixture's BMD-ED difference is larger.

3.0 TUTORIAL: STEP-BY-STEP GUIDE TO ANALYZING MIXTURES

3.1 Step 1: Setting Up the Tool

Make sure that macros are enabled in Excel either by clicking the "Enable Content" button at the top when opening Excel, or manually enabling all macros using the following steps:

1. Click the File tab
2. Click Options
3. Click Trust Center
4. Click Trust Center Settings
5. Click Macro Settings
6. Select "Enable all macros"

Note Before using MiST, users are encouraged to save an unused version of MiST with a different filename to ensure access to a backup copy if the user inadvertently deletes or disrupts codes hidden in various cells.

Note Users without administrative privileges may encounter difficulty enabling macros. MiST also includes a hidden Templates tab that defines data validated terms used in MiST. Users unable to enable macros or who wish to view the templates should consult with their employer's IT professionals to modify these settings. Users with administrative control can find additional guidance through Microsoft Excel Support.

Note Macros running other excel files may interfere with macros coded in MiST. Therefore, users should close any other macro-enabled excel file prior to opening MiST.

3.2 Step 2: Entering Mixture Properties

Mixture properties should be entered in the ‘Data Repository’ tab (Figure 3).

Note Blue cells in the ‘Data Repository’ and ‘Settings’ tabs can be edited by users.

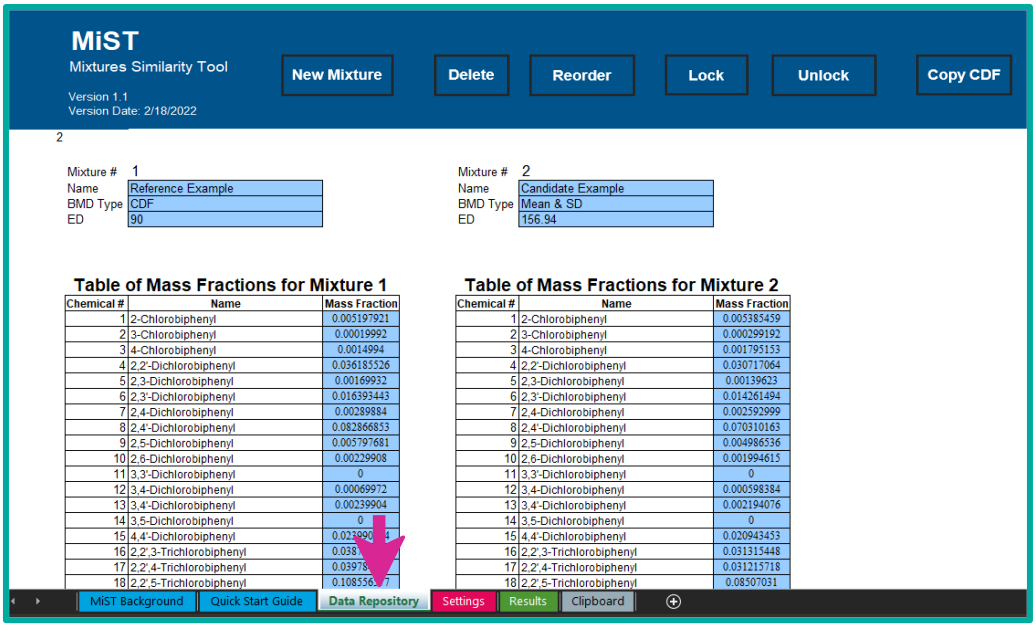

Figure 3. Data repository tab for entering mixture properties

3.2.1 Entering a New Mixture or Editing a Mixture

1. In the “Data Repository” tab, select a mixture to edit by clicking the "Unlock" button, or add another mixture using the “New Mixture” button (Figure 4).

2. Enter the name of the mixture in the “Name” field (Figure 6). The number of mixtures added will be shown in the top left corner (arrow).

Note In step 1, when prompted " Unlock Which Mixture #?", enter the desired mixture number (Figure 5). The mixture number automatically populates. The selected fields will turn blue when unlocked indicating that they can be edited. All three fields (i.e., Name, BMD Type and ED) must contain values. Once values are correctly entered, users must lock the mixture data in the Repository tab before running the analysis. To maximize the assisted entry in the tool, information should be entered for one mixture at a time before moving onto to data entry on subsequent mixtures.

Note In step 2, the mixture name in the Data Repository tab must be used in the 'Settings' tab to run the analysis. Mixture names are not case sensitive, but the spelling must agree across tabs.

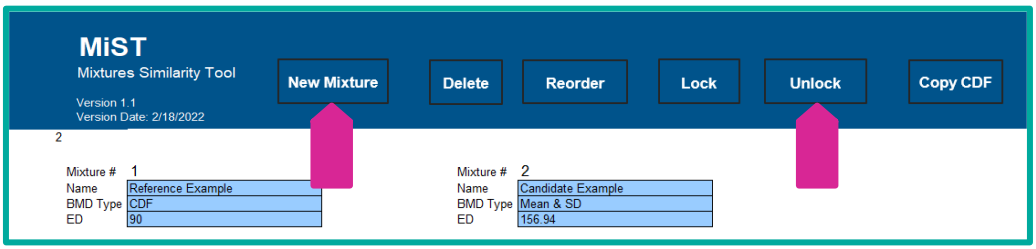

Figure 4. New mixture and unlock buttons

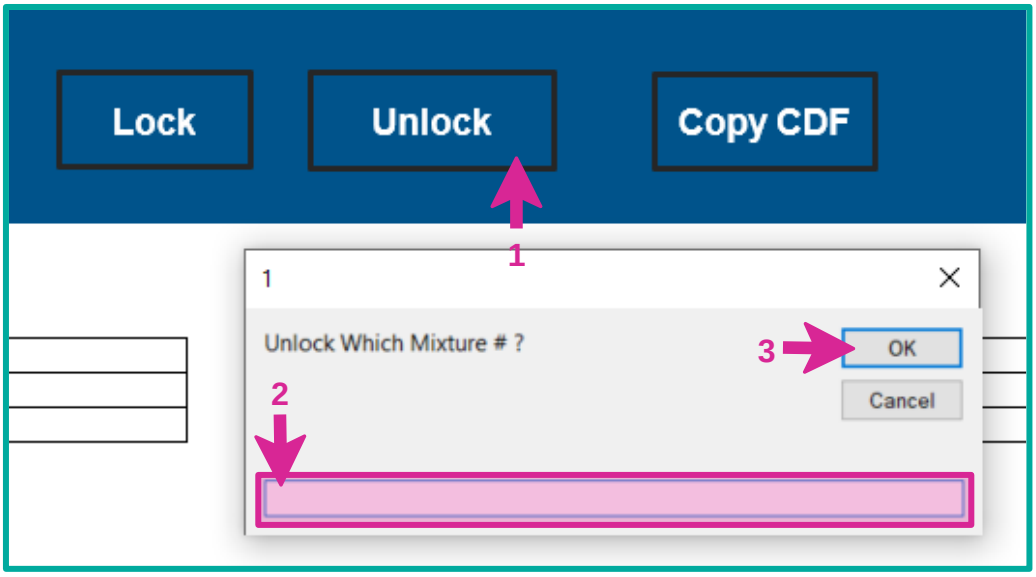

Figure 5. Specifying the mixture to unlock in the dialogue box

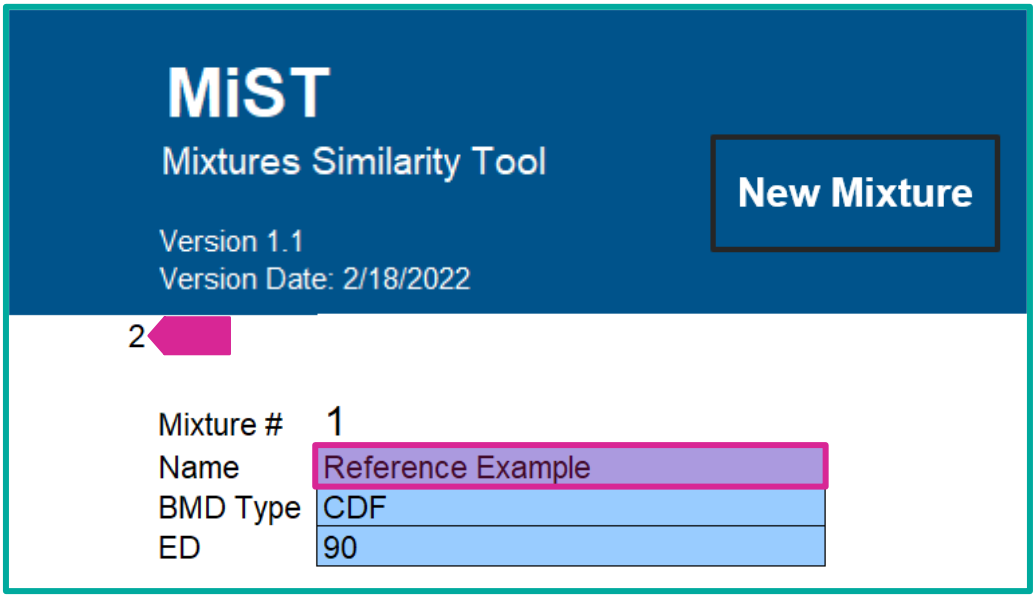

Figure 6. Entering the mixture name

3.2.2 Entering BMD Type

While the mixture is unlocked, select one of three BMD types (Figure 7):

- CDF,
- Mean & SD, or
- None.

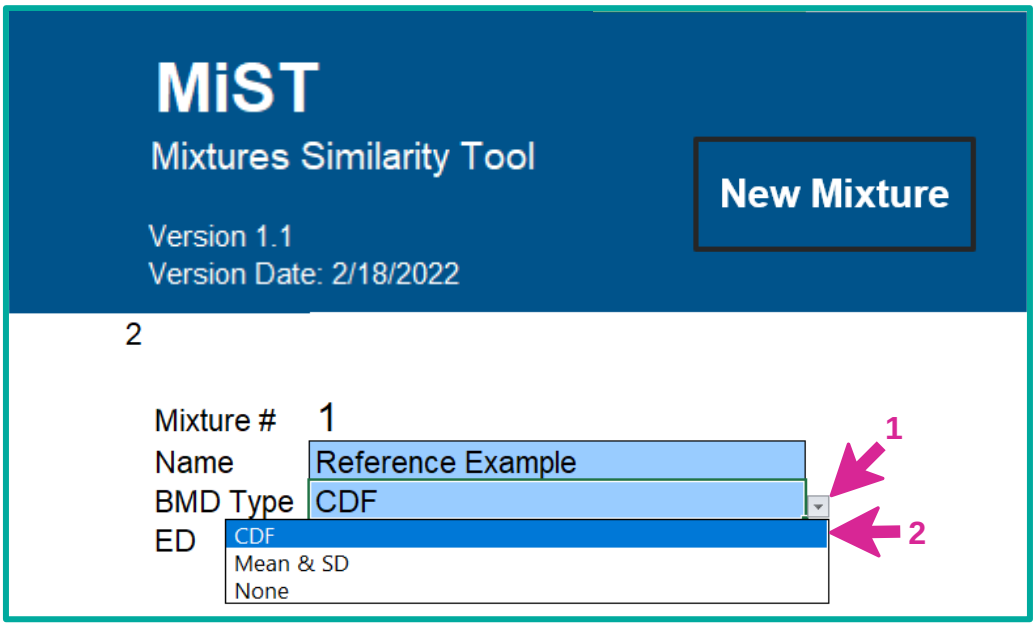

Figure 7. Selecting the BMD type

BMD Type: CDF

If a CDF is selected for this mixture (Figure 7), MiST will automatically move the cursor to the correct cell to enter CDF data (below the mass fraction data in the Data Repository tab). Values can be entered either as decimals or whole numbers. The user has two options to enter the data:

1. In the 'Data Repository' tab, manually enter the BMD values in relation to their corresponding percentile of its CDF.

2. In the 'Clipboard' tab, paste the available percentiles and BMD values in columns A and B, respectively (blue shaded columns). If all CDF values are not available, the user should add values for the percentiles that are available in consecutive lines in the clipboard. The "Copy CDF" and "fill" functions will assign values to the correct percentiles in the Data Repository tab.

    i. Return to the 'Data Repository' tab and click the "Copy CDF" button.

    ii. When prompted 'CDF for which mix?', enter the mixture number.

    iii. When prompted 'Is CDF data on clipboard?", select yes.

iv. The BMD values will then be pasted into the table below the mixture indicated above.

Note

Data output from BMDS are described elsewhere (U.S. EPA, 2020). Briefly, BMDS generates dataset-specific results workbooks containing separate worksheets for each model-option set combination that consist of tabular and graphical summaries of the modeling inputs and results. The BMD table from BMDS is found in the results workbook worksheets and can be copied and pasted into the MiST tool.

Note

The BMD values should be entered as a whole number or decimal greater than zero in non-decreasing order.

If a user does not have BMD values for every percentile, a piecewise linear interpolation function can be performed to estimate the values for the missing percentiles of the CDF. When manually imputing the BMD values, enter as many values as possible and select the 'Fill' button located next to the lower BMD table. Using the ‘Erase’ button here will delete any values in the BMD table. This will automatically occur if the "Copy CDF" button is used. See section 4.2 for more information on using the Fill and Erase buttons to interpolate BMD values as a CDF.

Only blue cells can be edited.

The clipboard allows users to store data for MiST runs. It also ensures all values are in an acceptable format for pasting into the Data Repository tab.

BMD Type: Mean & SD

Enter the mean and SD to be used as the distribution of the mixture’s BMD. Once “Mean & SD” is selected, enter the Mean and SD values at the bottom of that column (rows 229-232).

Note Once the BMD type is selected, MiST automatically moves the cursor to the cell where users will enter the mean and SD values.

BMD Type: None

Choosing “None” means that information on the BMD for this mixture will not be provided.

Note Mixtures with a BMD type of 'none' cannot be used as reference mixtures for the analysis

After selecting the BMD type, the tool will direct the user to input either the CDF or Parameters for the mixture. For more information on BMD types, see Section 4.2.

3.2.3 Entering the ED Value

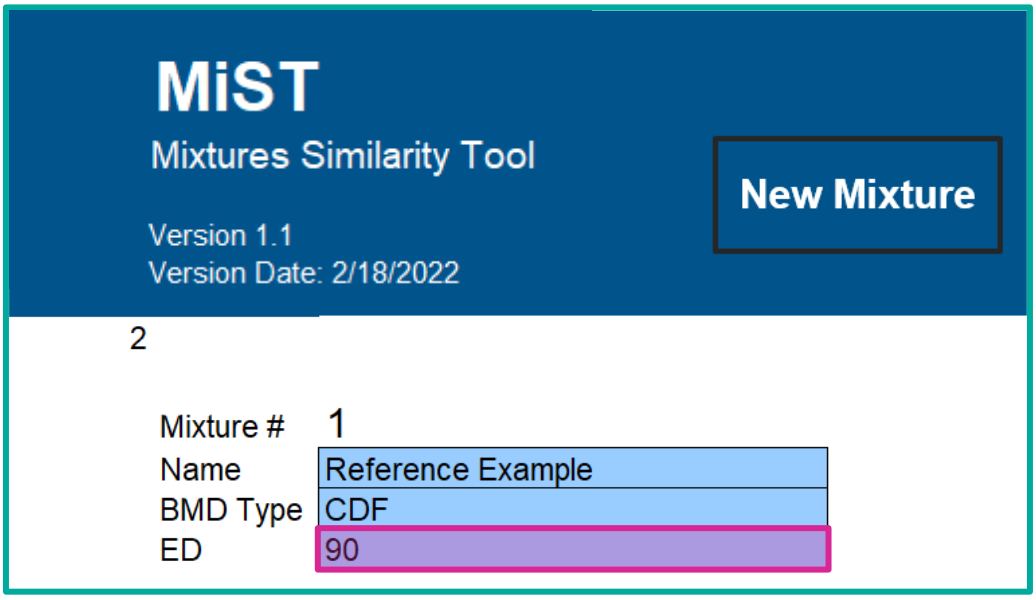

Figure 8. Entering the ED value in the ED field

Enter the ED value in the ED field.

Note For all reference mixtures and candidate mixtures in a data rich comparison[8], either an ED must be specified, or a CV must be entered on the Settings tab. The ED should be entered as a whole number or decimal greater than zero. It is used to generate the CV (calculated as the absolute value of the difference between the ED and BMD) if no CV is entered by the user on the Settings tab.

Note The preferred approach is to allow MiST to calculate a CV for each mixture based on mixture-specific EDs and BMDs, which is the default approach used in MiST if no CV is entered (See Section 1.2). The user has the option to assign an overriding CV (in the Settings tab) if needed but should not unless there is an appropriate justification for doing so.

[8] For data rich comparisons, the CV is the difference between the ED and BMD of the reference mixture or the difference between the ED and BMD of the candidate mixture, whichever is larger.

3.2.4 Entering the Mass Fractions

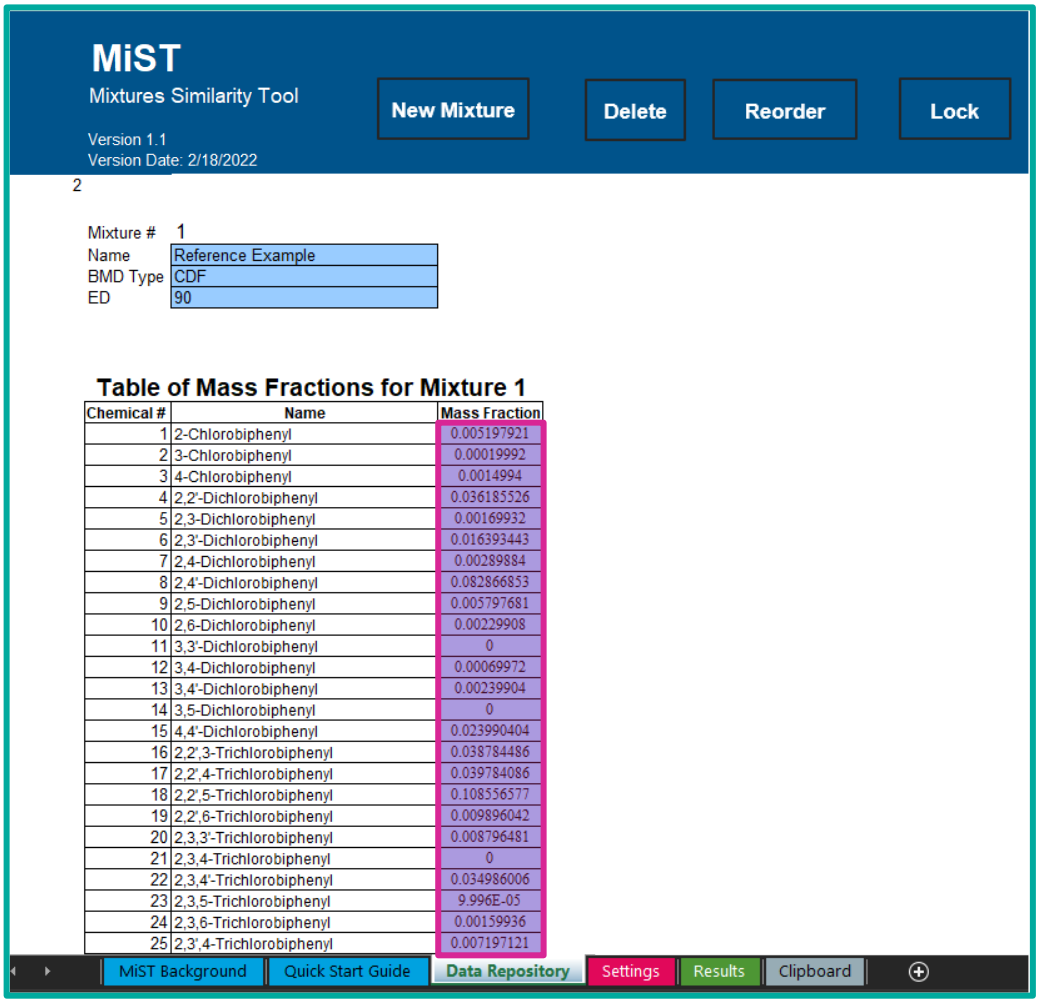

Figure 9. Enter mass fraction values

Note Enter the mass, concentration, or mass fraction values; if the total does not equal one (1), the tool will adjust the values so that they equal 1 when “Lock” is selected.

Note All mass fractions appear as zero when a new mixture is defined. Values must be entered by the user. If there are any missing values (i.e., blank cells), the user will receive an error message when attempting to lock the mixture.

3.2.5 Locking the Mixture

When the information is entered correctly for the mixture, select the "Lock" button at the top of the page (Figure 10) and enter the mixture number in the pop-up box. This will lock the mixture and make it available for analysis.

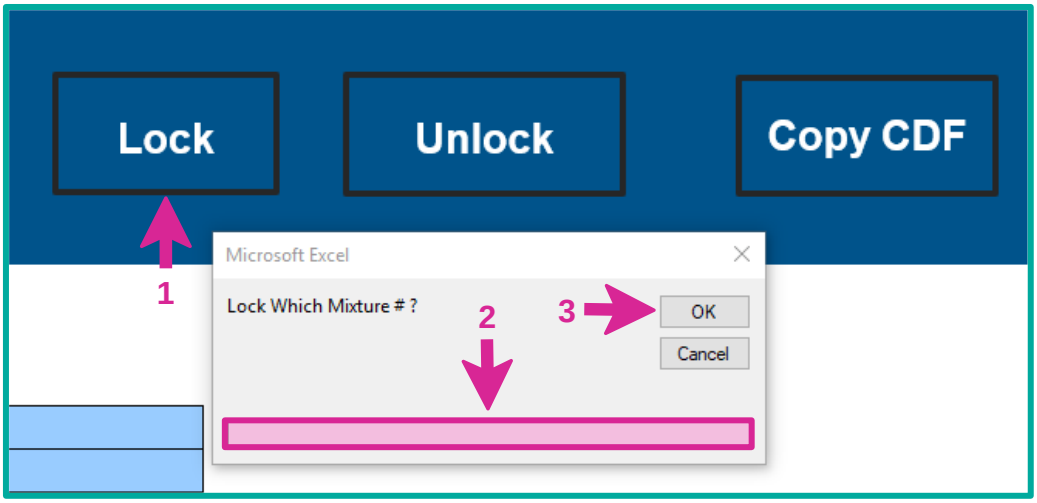

Figure 10. "Lock Which Mixture #?" dialogue box

3.2.6 Additional Tools in the "Data Repository" Tab

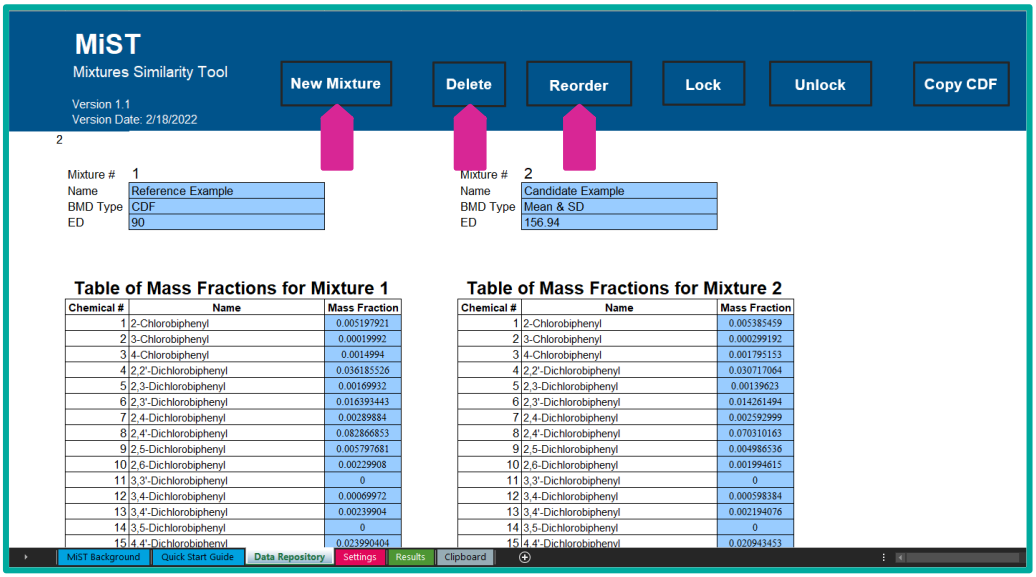

Figure 11. New mixture, delete, and reorder features

- **New Mixture** - generates a new, blank mixture entry (user can enter up to 99 mixtures)
- **Delete** - allows the user to delete a mixture entry; based on the mixture number
- **Reorder** - allows the user to reorder the mixture entries; the user is prompted to indicate the mixture to move by reporting the mixture number. The user will then be prompted to select the position to which the mixture should be moved.

3.3 Step 3: Determine the Parameters and Mixtures for Analysis

Enter parameters for the mixtures and the analysis in the Settings Tab. The user will

define the data inputs that define data rich or data poor run conditions and specify candidate and references mixtures.

3.3.1 Select Equal or Unequal Potencies

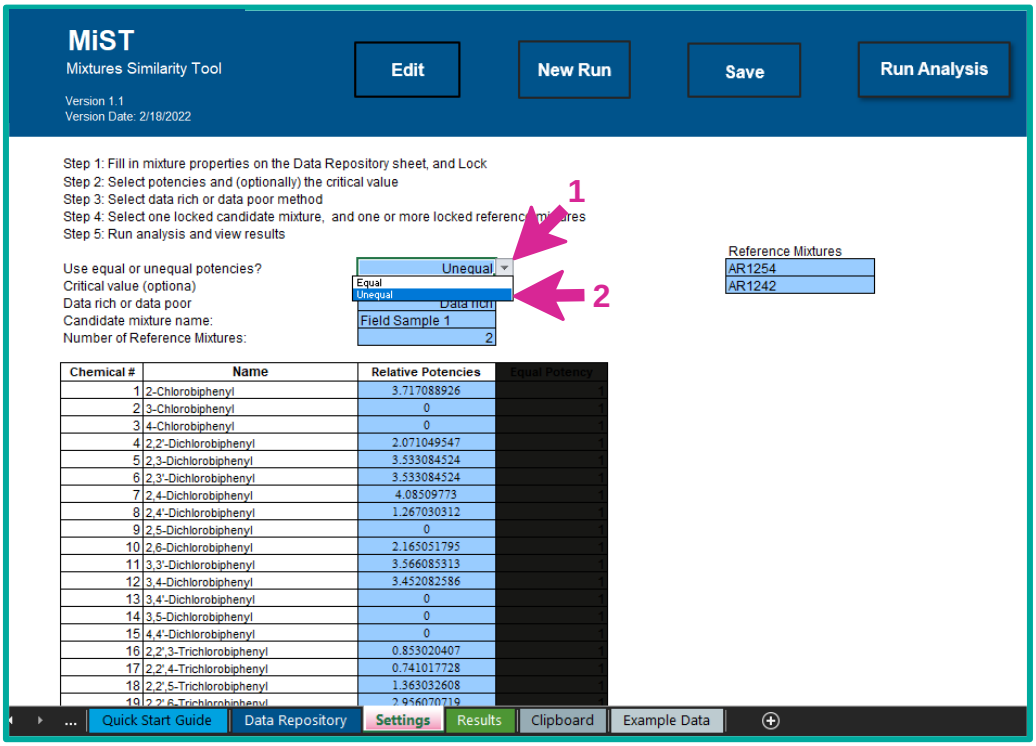

Figure 12. Selecting equal or unequal potencies

- **"Equal potencies"** signifies that the contributions of the mixture components are all treated equally. If using equal potencies, the relative potency estimates for each congener will be set to 1 in the column that appears below. Choose this option if the relative potencies of all mixture components (e.g., PCB congeners) are known to be equal or are not known. When congener potencies are assumed to be equal, congener contributions to the similarity determination are based only on mixture fractions.
- **"Unequal potencies"** signifies that a user will input congener-specific relative toxicological potencies. If using "unequal potencies", the user must assign a relative toxicological potency estimate to each PCB congener in the blue cells in the "Relative Potencies" column. When using this option, each congener's mixture fraction and its resulting contribution to the similarity determination is scaled based on congener-specific potency. Any congener potencies left blank will be assumed to be zero. Congener potencies assigned a zero value will be automatically populated in the results tab after the user initiates the run.

Note: The relative toxicological potency estimates entered in MiST are derived by users prior to conducting a mixtures similarity analysis or are available in the published literature (e.g., Holland and Pessah, 2021; U.S. EPA, 2010). Toxicological potency for each congener is estimated relative to an index congener, which is typically the congener with the highest known potency and

is assigned a relative potency estimate of 1. Once relative toxicological potency estimates are entered in this format and the "run analysis" action is initiated, MiST will autoscale the potencies so that the sum equals 209, the maximum number of components possible in a PCB mixture. This action makes the results of the equal and unequal potency analyses more directly comparable (see Table 4 for more information).

Note The same congener-specific potency estimates are used for all mixtures in an analysis. Additional information on relative toxicological potency is presented in section 1.2.

3.3.2 Enter the Critical Value

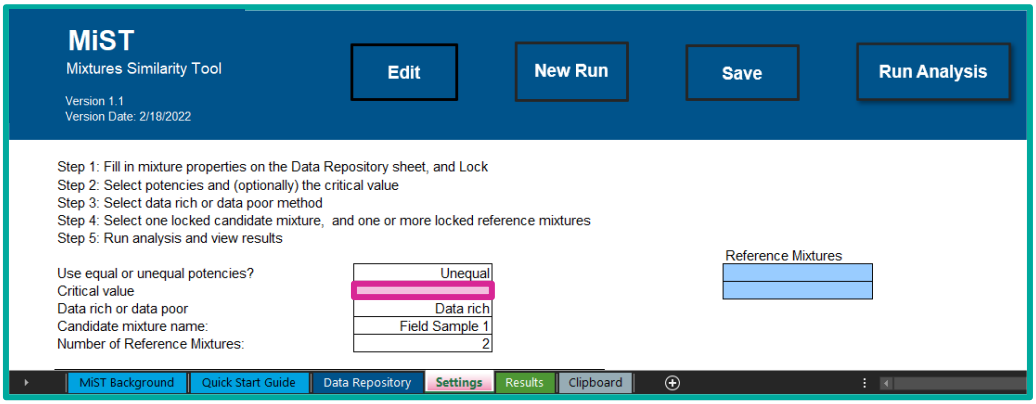

Figure 13. Entering the CV

The optional user-specified CV should be entered as a whole number or decimal greater than zero.

Note Selection of an appropriate CV requires expert judgment (see Section 1.2). If no value is entered, the CV will be determined by the default (and preferred) approach of calculating the absolute difference between the user-specified BMD and the ED, required mixture-specific entries in the Data Repository tab. Users should only override the default approach with an appropriate justification.

3.3.3 Select Data Rich or Data Poor

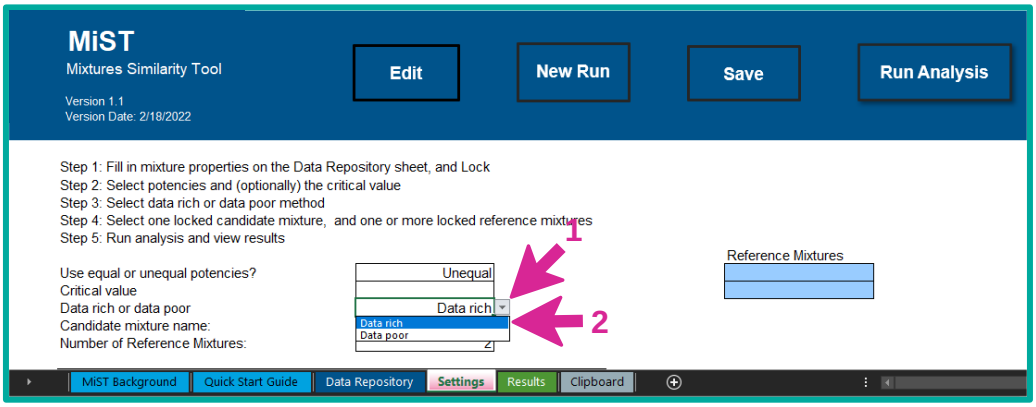

Figure 14. Defining inputs as data rich or data poor

- **Data rich** indicates that the BMD for the candidate mixture is available.
- **Data poor** indicates that the BMD for the candidate mixture is NOT available.

3.3.4 Enter the Candidate Mixture Name

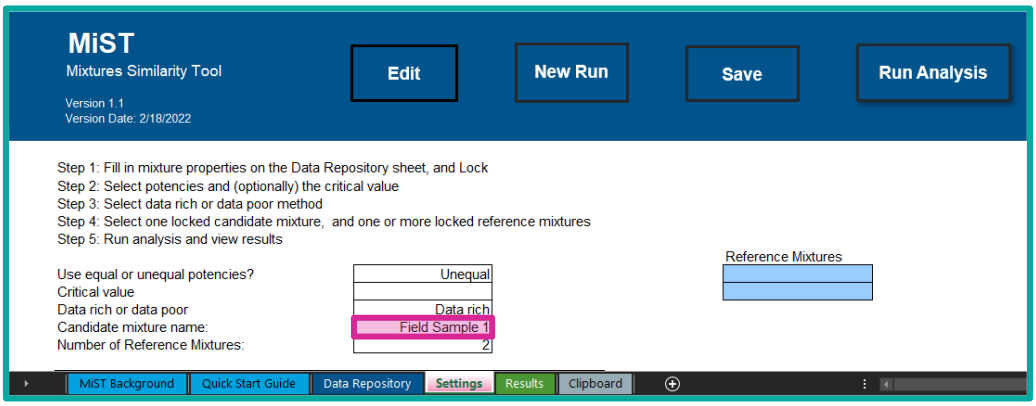

Figure 15. Entering candidate mixture names

Determine which mixture from the 'Data Repository' tab will act as the candidate mixture for the analysis and enter it in the “Candidate mixture name” field.

Note Although mixture names are not case sensitive, users should enter the name for this mixture in full as spelled in the 'Data Repository' tab.

3.3.5 Enter the Number of Reference Mixtures

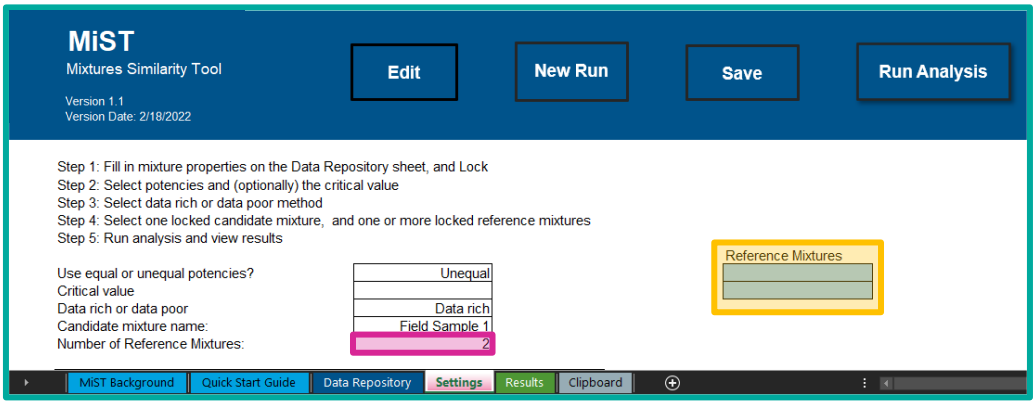

Figure 16. Entering number of reference mixtures

Enter a positive, whole number of reference mixture(s) to compare to the candidate mixture. The appropriate number of boxes will pop up in column H.

Note Each reference mixture must have a BMD entered in the 'Data Repository' tab.

3.3.6 Enter the Reference Mixture Names

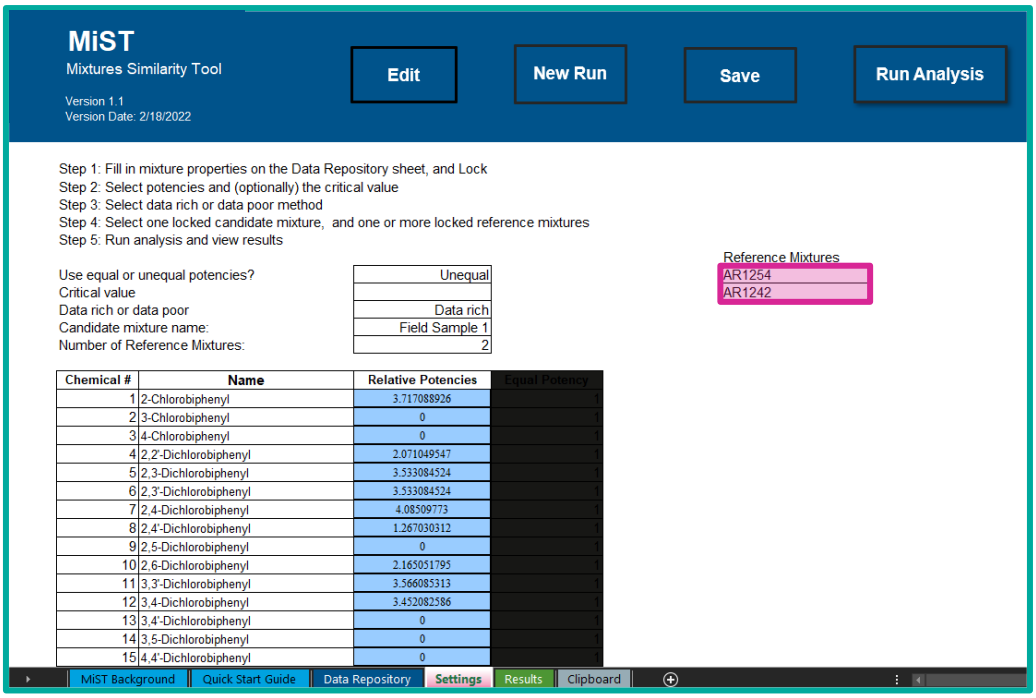

Figure 17. Entering reference mixture names

1. Enter the number of Reference Mixtures in the field. This will generate the Reference Mixtures fields to the right.

2. Determine which mixtures from the 'Data Repository' tab will act as the reference mixture(s) for the analysis. Enter the name for each mixture as it is written in the 'Data Repository' tab in the box to the right

3.3.7 Run the Analysis

When the information is entered correctly for the analysis, select the "Run Analysis" button at the top of the page. A successful analysis will result in a popup message that says, "Analysis Completed."

Note All mixtures must be locked on the Data Repository Tab before the analysis can be run.

3.3.8 Additional Tools in the Settings Tab

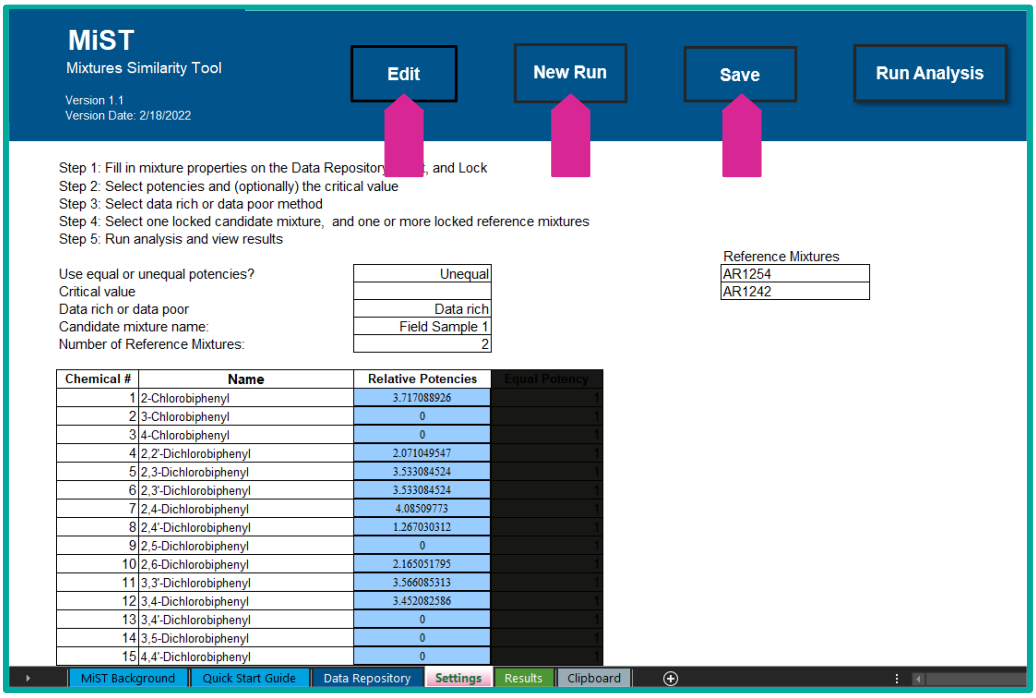

Figure 18. Additional tools in the 'Settings' tab

- **Edit** – Unlocks the information for editing but does not erase it.
- **New Run** – Clears the information from the 'Settings' tab; the 'Data Repository' tab will not be changed
- **Save** – Allows the user to save the workbook; it is recommended to save each analysis under a new filename. This is the same as the Excel "Save As" feature.

3.4 Step 4: Review the Results

These properties/actions are housed in/can be accomplished in the 'Results' tab.

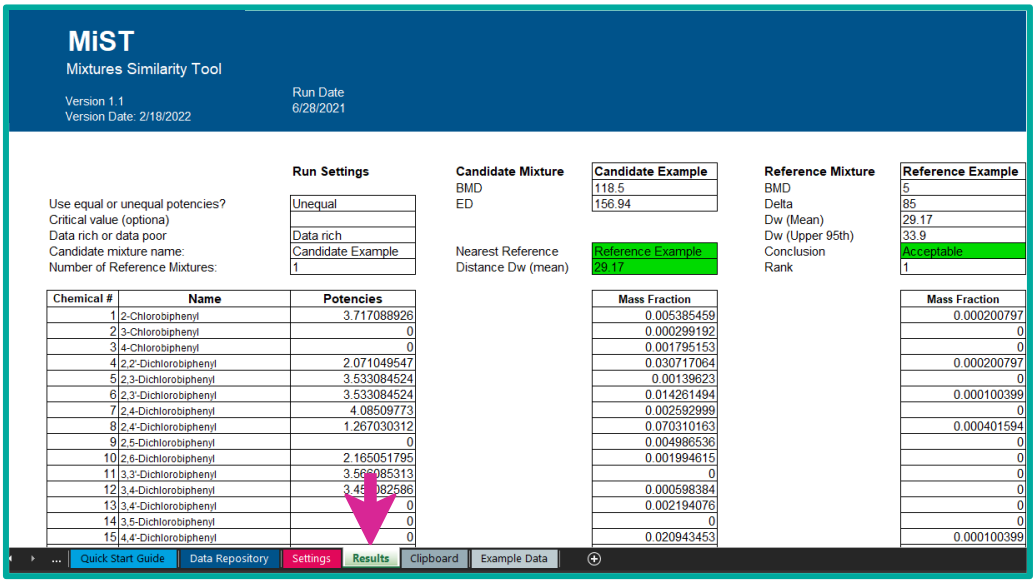

Figure 19. Reviewing the results on the 'Results' tab

3.4.1 Tool Run Settings

The settings selected for the analysis are returned at the top of the Results Tab.

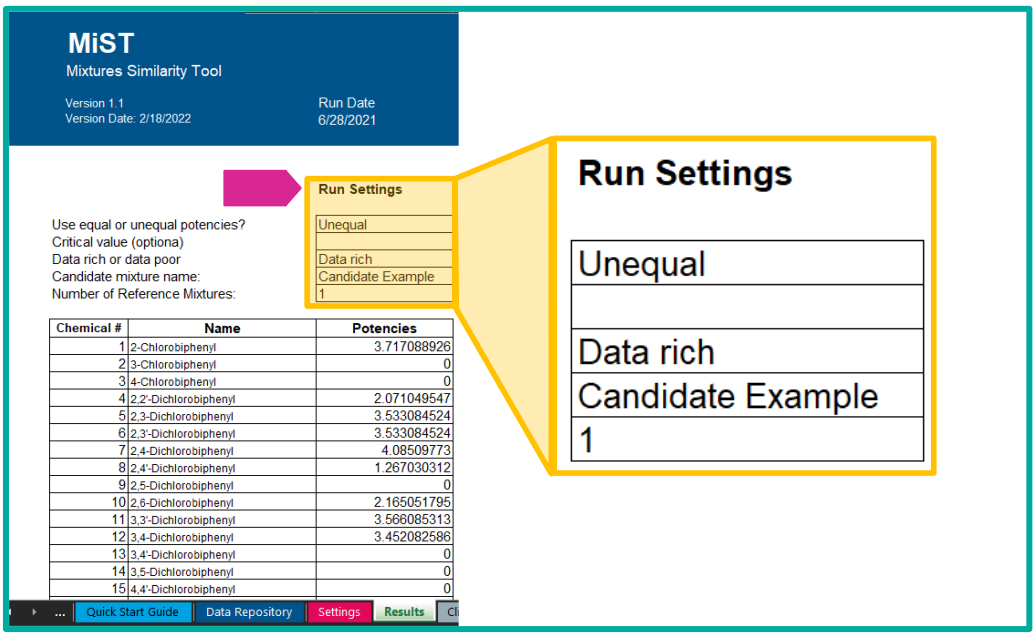

Figure 20. Run Settings entered on the 'Settings' tab populate on the 'Results' tab

3.4.2 Results for the Candidate Mixture

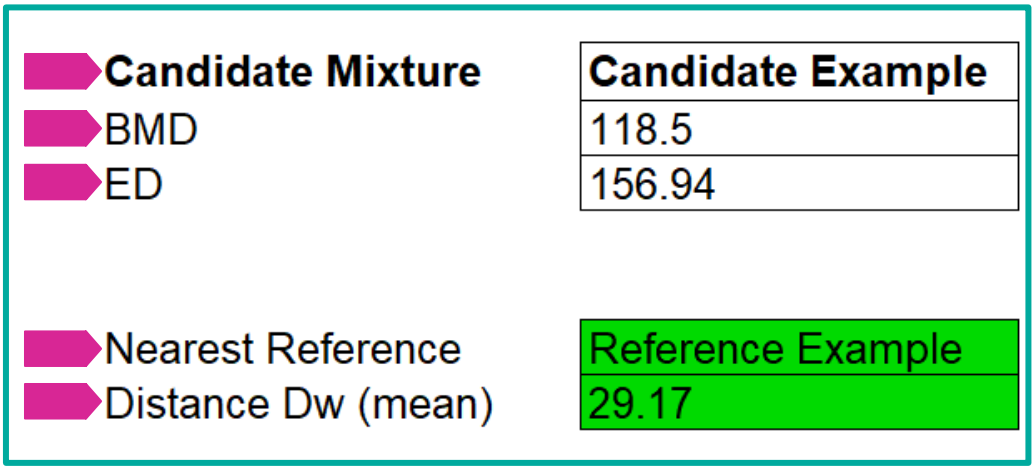

Figure 21. Candidate mixture results for BMD, ED, nearest reference, and distance Dw (mean)

- **Candidate Mixture** – The candidate mixture name is returned.
- **BMD** – If mean and SD were entered, the mean BMD is returned. If a CDF was entered, the 50th percentile is returned.
- **ED** – The ED value of the candidate mixture is shown (data rich runs only).
- **Nearest Reference** – Indicates the name of the reference mixture that is the most similar to the candidate mixture as determined by calculating and comparing the Euclidean distance between the candidate mixture and each reference mixture.
- **Distance *Dw* (mean)** – The average weighted Euclidean distance between the nearest reference mixture (indicated above) and the candidate mixture.

Note The Nearest Reference and Distance ***Dw*** (mean) boxes are colored green if the nearest reference is deemed “Acceptable” or are colored red otherwise.

3.4.3 Results for the Reference Mixture

Reference Mixture	Reference Example
BMD	5
ED	90
Delta	85
Dw (Mean)	22
Dw (Upper 95th)	25.798
Conclusion	Acceptable
Rank	1

Figure 22. Reference mixture results for BMD, delta, Dw (mean), Dw (upper 95th), conclusion, and rank

- **Reference Mixture** – The reference mixture name is returned.
- **BMD** – If mean and SD were entered, the mean BMD is returned. If a CDF was entered, the 50th percentile is returned.
- **ED** – The ED for the reference mixture is returned, if specified by the user.
- **Delta** – The CV for the analysis, either from the Settings tab or calculated as the absolute value of the difference between ED and BMD for that reference.
- ***Dw* (Mean)** – The average Euclidean distance between the reference mixture and the candidate mixture.
- ***Dw* (Upper 95th)** – The 95th percentile of the Euclidean distance between the reference mixture and the candidate mixture.

Conclusions:

- **Acceptable** – The ***Dw*** (Upper 95th) value for the reference mixture is less than the CV (Delta); consider the reference mixture SIMILAR to the candidate mixture.
- **Not Acceptable** – The ***Dw*** (Upper 95th) value for the reference mixture is greater than the CV (Delta); do NOT consider the reference mixture sufficiently similar to the candidate mixture.

Note Rank: The reference mixture with the smallest ***Dw*** (Mean) is the considered the best match and assigned the rank of 1.

4.0 MIST DOCUMENTATION

4.1 Assumptions

The main assumptions built into the current tool are:

1. The chemical mixtures consist of PCB congeners.
2. All congener mass fractions are specified without any estimate of their potential uncertainty or error.
3. The mass fractions sum to one, meaning no chemicals other than PCBs are present in the mixture.
4. The relative potency of each congener is specified without an estimate of uncertainty or error.
5. Each reference mixture has a BMD, or BMD-like estimate, with an associated uncertainty distribution, for some health effect.
6. The health effects for all reference mixtures and any candidate mixtures in a single analysis are comparable.

Note It is the responsibility of the user to assure that BMDs refer to comparable health effects.

7. The "relative toxicological potency" of each congener is an estimate of the congener's potency relative to the potency of an index congener, which is typically the congener with the highest known potency and is assigned a relative potency estimate of 1.

Note The BMD of a mixture is the weighted sum of the BMD of its congeners, there is no synergism or antagonism among congeners (dose additivity).

8. The BMD distribution for the candidate mixture is independent of those for the reference mixtures, meaning the covariance between the two distributions is assumed to be zero.

Note Marshall et al. (2013) allowed for the use of variance-covariance matrices, but this is not currently supported in MIST.

9. The difference between the candidate and reference BMDs (calculated as the weighted Euclidean distance in equation 1, section 3.3) defines how toxicologically "similar" the two mixtures are to each other.

4.2 Characterizing the Uncertainty in Reference BMD

There is some degree of uncertainty in the BMD for each mixture. Two methods are available in MiST for describing this uncertainty: CDF or mean with SD.

If the CDF method is used, the user specifies any integer percentiles they may have for BMD values on from the 1st to 99th, and optionally the minimum and/or the maximum as well. It is assumed that any non-specified percentiles are uniformly spaced between those that are given. The tool has a "Fill" button that performs this interpolation. It is necessary that this is done before comparing a candidate mixture to reference mixtures.

If the mean and SD are specified, it is assumed that the distribution is normal. The tool then generates the integer percentiles from that normal distribution automatically. It also sets the maximum and minimum at plus or minus three SDs from the mean, as these points are a little beyond the 99th and first percentiles, respectively.

Once all the percentiles along with the minimum and maximum are determined, the tool may start the Monte Carlo iterations for calculating distance between the candidate mixture and each reference mixture.

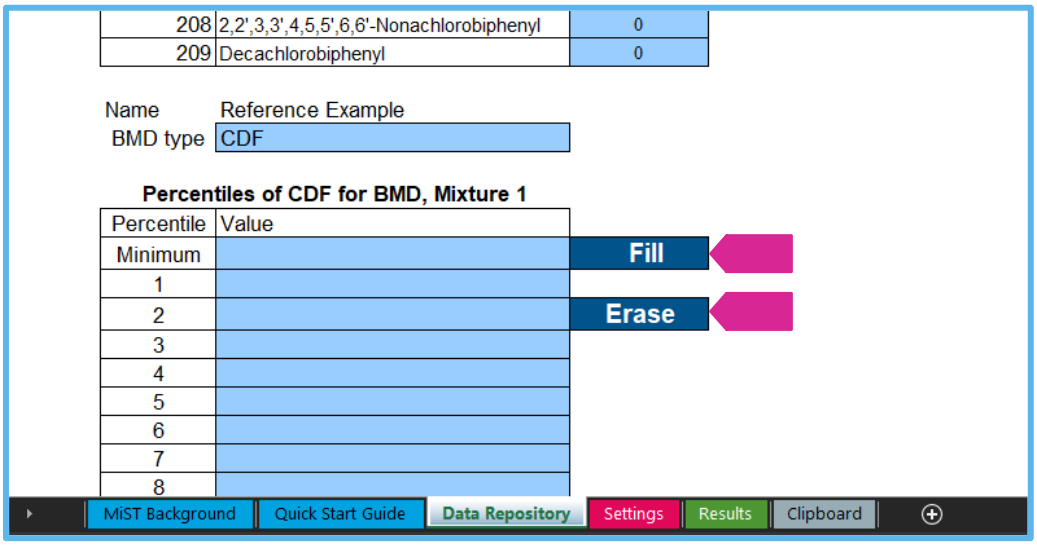

Figure 23. Fill and erase buttons for interpolating or clearing CDF values

Fill and Erase buttons are available to the user to aid in entering or clearing BMD values entered as a CDF in the Repository Tab. The "Erase" button will remove all values and formats entered into the CDF table. The Fill and Erase buttons are only available if the mixture is unlocked. The Fill button can be used to interpolate a partial list of CDF percentiles using the following rules:

1. The CDF is not complete until all 101 values are specified. These are the minimum, percentiles 1-99, and the maximum.

2. BMD values can only be entered for the whole number percentiles (1-99) listed. BMD values may be entered as decimals or whole numbers.

3. No value may be negative.

4. No value may be smaller than a value at a lower percentile.

5. If two or more values are entered, the user may use the “Fill” button to populate the rest.

Note The “Fill” button uses the following logic:

Each integer percentile gap (between two specified values) is filled by linear interpolation. Interpolation can be used to fill towards the minimum value, the maximum value, or in both directions.

Values above the highest specified percentile are filled using the same slope (increase per percentile) as occurs between the two largest specified values. The same slope (decrease per percentile) applies to interpolation to the minimum.

4.3 The Monte Carlo Approach

The concept behind MiST is that the difference in BMD between the candidate and a reference mixture can be characterized by randomly sampling each uncertainty distribution and calculating the distance between the mixtures. On one iteration, the distance is defined as

$$\boldsymbol{Dw} = \sqrt{\sum_{j=1}^{209} \boldsymbol{P_j}\left(\boldsymbol{T_r}\boldsymbol{\alpha_{jr}} - \boldsymbol{T_c}\boldsymbol{\alpha_{jc}}\right)^2} \qquad (1)$$

- $\boldsymbol{j}$ label uniquely identifying each of the 209 PCB congeners.
- $\boldsymbol{P_j}$ potency of congener $\boldsymbol{j}$, relative to the index potency of one
- $\boldsymbol{T}$ estimate of BMD as sampled from appropriate distribution (either $\boldsymbol{r}$ or $\boldsymbol{c}$)
- α_j mass fraction of congener $\boldsymbol{j}$ in the mixture (either $\boldsymbol{r}$ or $\boldsymbol{c}$)
- $\boldsymbol{r}$ reference mixture
- $\boldsymbol{c}$ candidate mixture

If $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$ were known precisely, the distance in equation (1) between the candidate and each reference mixture would just be evaluated once. The point is that both $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$ have uncertainty, so their true values are not known. The Monte Carlo approach involves selecting random samples for $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$ and evaluating the distance for that pair of values. The MiST tool currently samples $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$ independently, under the assumption that there is no appreciable correlation expected between them. This produces a set of $\boldsymbol{Dw}$ values, one per iteration. Ten thousand iterations are run for each analysis.

It may be noted that even if the candidate mixture had the same BMD distribution as some reference mixture, the distance between them would never be zero unless $\left(\boldsymbol{T_r}\boldsymbol{\alpha_{jr}} - \boldsymbol{T_i}\boldsymbol{\alpha_{jc}}\right)^2$ were zero for all congeners $\boldsymbol{j}$. This would require that the same sample values be drawn for $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$ on the same iteration and that all the mass fractions were the same in the two mixtures, which is unlikely. Apart from this, all $\boldsymbol{Dw}$ are positive.

The method requires that the "true" $\boldsymbol{Dw}$ should be below the CV threshold of acceptability for the candidate to be considered "similar" to a given reference. Since the true value of $\boldsymbol{Dw}$ remains unknown, it is required that 95% of the Monte Carlo iterations produce a sample $\boldsymbol{Dw}$ below the CV threshold for acceptance. Since all iterations are equally valid estimates, this equates to a 95% confidence that the true $\boldsymbol{Dw}$ is below the CV.

4.4 The Data Rich Case

All reference mixtures must have a BMD distribution specified. This is optional for the candidate mixture. If a candidate uncertainty distribution is given, then it is the "data rich" case, discussed here. When the distribution for the candidate is not known, it is called the "data poor" case and is discussed in the next section.

The data rich case uses equation (1) with two independently drawn random samples for $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$. In practice, this is achieved by first generating random samples from a uniform distribution between zero and one. These are then rounded to two decimal places and multiplied by 100. The corresponding percentile from the distribution is then used. For both cases (either the CDF or the mean with SD), all percentiles from 1 to 99 are computed and stored, along with a minimum and maximum. If the uniform random sample is below 0.005, it rounds to zero to two decimal places, and the minimum is used. If it is above 0.995 the maximum is used. Each of these occurs 0.5% of the time. Each percentile from 1 to 99 occurs (on average) 1% of the time. The MiST tool is currently set to perform 10,000 iterations per run, so each percentile will be selected 100 times on average. The values of the other variables in equation (1) do not change on each iteration. This method produces 10,000 estimates of $\boldsymbol{Dw}$ between the candidate and the given reference.

If there is more than one reference mixture being compared to the candidate, the above steps are repeated for each reference. In practice this means another 10,000 random samples are generated for each reference. This is relatively fast and avoids any possibility of the results for one reference influencing the results for another.

Preliminary tests demonstrated that 1,500 iterations should be sufficient for convergence (Hong et al., 2016). However, that conclusion depends in part on the specific uncertainty distributions being used, so 10,000 iterations were selected to reliably ensure convergence.

Once all the iterations are complete, the mean and 95th percentiles of the $\boldsymbol{Dw}$ samples are found and reported on the Results tab for each reference. If the 95th percentile is below the threshold, the "Conclusion" cell for that reference is colored green. If not, it is colored red. Under the candidate mixture on the Results tab, the reference with the smallest mean $\boldsymbol{Dw}$ is indicated and colored green if appropriate. MiST ranks references according to which has the closest 95th percentile of $\boldsymbol{Dw}$.

4.5 The Data Poor Case

The data poor case uses equation (1), only the distribution for $\boldsymbol{T_c}$ is assumed to be the same as for $\boldsymbol{T_r}$. Even so, the Monte Carlo approach draws separate samples for $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$. This ensures that if the candidate happens to have a BMD distribution that essentially duplicates that of $\boldsymbol{T_r}$, the results will still be independent under assumption (9). If the same samples were used for $\boldsymbol{T_r}$ and $\boldsymbol{T_c}$, that would correspond to an assumption of a covariance of unity, contrary to assumption (9).

If the candidate is being compared to several references, then the sampled $\boldsymbol{T_c}$ values are different for each reference. This is already the case using the data rich method because different random percentiles are drawn on each iteration for each reference. In the data poor case, the values corresponding to the same percentiles of $\boldsymbol{T_c}$ are also different, depending on which reference it is being compared with. As a practical matter, the tool does not distinguish the data rich and data poor cases, apart from the matter of where it looks up the percentiles for $\boldsymbol{T_c}$.

Once the iterations are complete, as in the data rich case the mean and 95th percentile of the $\boldsymbol{Dw}$ samples are found and reported on the Results tab for each reference. If the 95th percentile is below the threshold, the “Conclusion” cell for that reference is colored green. If not, it is colored red. Under the candidate mixture on the Results tab, the reference with the smallest mean $\boldsymbol{Dw}$ is indicated and colored green if appropriate.

5.0 MIST VALIDATION

Validation of MiST included review by internal and external reviewers. Original reports detailing numerical results and codes used for validation tests from internal and external reviewers are summarized here and are available upon request (Gennings, 2017). Internal validation was conducted by ICF and submitted to EPA is also available upon request (Hong et al., 2016).

A. The following tests were conducted to validate underlying assumptions of MiST, and results are summarized in Table 6.

- Run the modified test case from Marshall et al. (2013). NOTE: the test case in Marshall et al. (2013) was an analysis of pyrethroids that included fewer mixture constituents than the 209 supported by MiST. Thus, there were some differences between this analysis and the specific analyses presented in that publication due to the nature of the test example.
- Run the same mixture as both candidate and reference
- Use the same BMD distribution for candidate and reference but alter the mass fractions substantially

Table 6. Summary of Internal MiST Validation

Scenario	Tool Settings	Expected Result	Confirmation
Compare two chemicals with the exact same mixing ratios, BMD, SE of BMD, and ED	• data rich, equal potency • data rich, unequal potency • data poor equal potency • data poor, unequal potency	Distance (***Dw***) should be 0 and SE should be 0 in the data poor case	Expected results confirmed
Compare two chemicals with the exact same mixing ratios, BMD, and ED, but with a different SE of BMD	• data rich, equal potency • data rich, unequal potency	Distance (***Dw***) should be 0 but SE of distance should increase with increasing SE of BMD	Expected results confirmed
Compare two chemicals with the exact same mixing ratio and SE of the BMD but different BMD/ED	• data rich, equal potency • data rich, unequal potency	The distance and SE (distance) should increase for increasing difference between the BMDs	Expected results confirmed
Compare two chemicals with the exact same BMD/ED and SE of the BMD but different mixing ratios	• data rich, equal potency • data rich, unequal potency	Distances will likely be maximized across all test scenarios	Expected results confirmed

B. ICF validated the implementation of the Monte Carlo method to measure standard error of the BMD by conducting the following tests (Trgovcich et al., 2021):

- Compare the Monte Carlo results to the earlier Delta method that was recommended in the Marshall and colleagues.
- Test the assumption of independence between the two mixtures used in the Monte Carlo and delta method tests by comparing the standard error of the distance estimates when the variance-covariance matrix was used as the data prescribe versus when the variance-covariance matrix was forced to 0 (the independent assumption case).

These validation tests confirmed that the two methods provide similar results, with the difference being larger in the data rich case (which uses two different BMD distributions) than in the data poor case (which assumes both mixtures have the same BMD distribution). With respect to the assumption of independence between the mixtures, the effect of using the Monte Carlo method compared with the delta method gives a larger proportional difference in these variable values than the assumption of independence alone. Taken together, results of these tests confirmed that the two methods are in sufficient agreement to justify using the Monte Carlo method.

Additional validation tests were conducted to validate MiST features added or modified during MiST development **(**Trgovcich et al., 2021**).** General function tests were also conducted and confirmed that features added or modified during development did not alter the MiST code or expected conclusions.

C. External validation of MiST was conducted by the academic research group headed by Dr. Chris Gennings in 2017. Dr. Gennings' report under EPA Contract No. EP-C-14-001 Work Assignment 3-15 was entitled "Expert Assistance for the IRIS Draft Toxicological Review of PCBs - Effects other than Cancer: Testing of Mixtures Similarity Tool" (Gennings, 2017). External reviewers conducted the following analyses:

- Compared 4 Aroclor mixtures (AR1254, AR1242, AR1016, and AR1260) of PCBs using specific test conditions, including data rich, data poor, weighted, and unweighted analyses. Following the published method (Marshall et al., 2013), researchers developed and used a SAS code to test for sufficient similarity between PCB mixture AR1254 (reference mixture) and each of the three candidate mixtures: AR1242, AR1016, and AR1260. External validation tests confirmed that MiST generated expected results for these tests.
- The preferred statements for understanding the results of MiST were reviewed and revised are as follows:
  - Candidate mixture is similar to reference mixture: claimed when the null hypothesis is rejected, i.e., when the upper 95% confidence limit is less than the similarity boundary.
  - Unable to claim similarity between the candidate and reference mixtures: stated when the null hypothesis is not rejected, i.e., when the upper 95% confidence limit exceeds the similarity boundary.
- Reviewed the Monte-Carlo simulation for determination of the estimated standard error of the distance ($\boldsymbol{Dw}$) and recommended for a minimum number of iterations to be performed for the Monte Carlo simulation that has a meaningful impact on tool conclusions and analysis. Tests concluded that 1500 iterations should be sufficient. The revised tool now builds in a safety factor by using 10,000 iterations.

Note Following internal and external validation, several small modifications were made to the MiST to improve efficiency and resolve identified issues with layout and presentation of the results. A summary of these changes are available upon request (Trgovcich et al., 2021).

6.0 TROUBLESHOOTING

1. **Saving files.** Before using MiST, save a backup version of MiST with a different filename such as MiST 1.0_backup.

Note Once data are entered into the tool and saved, Microsoft Excel will overwrite the original file. Saving a pristine backup file will ensure the user has access to the original tool.

2. Macros running other excel files may interfere with macros coded in MiST. Therefore, users should close any other macro-enabled excel file prior to opening MiST.

3. **Entering CDF or Mean and SD data.** It may be necessary to scroll past the mass fraction data to enter the CDF or mean and SD data.

4. **Problems accessing the “undo” function in Excel.** Microsoft Excel disables the "undo" feature after running macros. Macros are executed when selecting any button at the top of the Settings and Data Repository tabs. Users are encouraged to save their work before overwriting existing data.

5. **Problems pasting data.** VBA scripts may be activated upon selecting cells which may preclude pasting data copied before selecting the correct cell. To avoid problems pasting data, users should first select the cell where data will be pasted, then copy the data, then paste.

7.0 GLOSSARY

Benchmark Concentration (BMC) or Benchmark Dose (BMD)	An exposure due to a dose of a substance associated with a specified low incidence of risk, generally in the range of 1% to 10%, of a health effect; or the dose associated with a specified measure or change of a biological effect. BMCs/BMDs can be expressed in a variety of units, including mg/kg-d and mg/kg as well as ppm and mg/m^3. A benchmark expressed as a concentration (e.g., ppm) is generally referred to as a BMC while a benchmark expressed as a dose (e.g., mg/kg-d) is a BMD. BMDs/BMCs are functionally equivalent for the purpose of MiST, but it is essential to match the units of the BMDs/BMCs compared in a data rich analysis.
Benchmark Dose Software (BMDS)	Software developed by EPA to estimate reference doses (RfDs), reference concentrations (RfCs), and slope factors, which are used along with other scientific information to set standards that protect human health from the effects of chemical exposures.
Candidate mixture	Mixtures that have been selected for risk evaluation but lack adequate dose-response data for estimating levels of exposure associated with adverse effects. In many cases, these represent mixtures encountered in the environment.
CatReg Software	Software developed by EPA to perform categorical regression analyses on toxicity data after effects have been assigned to ordinal severity categories (e.g., no effect, adverse effect, severe effect) and associated with up to two independent variables corresponding to the exposure conditions (e.g., concentration and duration) under which the effects occurred. CatReg calculates the probabilities of the different severity categories over the continuum of the variables describing exposure conditions.
Critical Value (CV)	The value against which the 95^{th} percentile distance is compared to determine sufficient similarity. The preferred method is to calculate this value for each reference mixture as the difference between its BMD and a mixture-specific ED value. Alternatively, a single generic CV that is applied to all reference mixtures can be set by MiST users but only with appropriate justification.

Cumulative Distribution Function (CDF)	A function that gives the probability that a random variable is less than or equal to the independent variable of the function.
Dw	Weighted Euclidean distance between reference mixture and candidate mixture
Effective Concentration (EC) or Effective Dose (ED)	Effective concentration or dose associated with a biological effect in x% of the individuals. Selection of EC/ED values for use in mixtures similarity analyses requires careful consideration and is described further in Section 1.2. Example units for an EC include ppm and mg/m^3; example units for an ED include mg/kg-day.
Extra Risk Concentration (ERC) or Extra Risk Dose (ERD)	An exposure due to a dose of a substance associated with a specified low incidence of risk, generally in the range of 1% to 10%, of a health effect; or the dose associated with a specified measure or change of a biological effect. Example units for an ERC include ppm and mg/m^3; example units for an ERD include mg/kg-day.
Index Congener	The congener selected as the basis for standardization of toxicity of components in a PCB mixture. The index congener must have a clearly defined dose-response relationship.
Mass Fraction	Ratio of the mass of a component in the mixture to the total mass of the mixture. Mass fraction is dimensionless. For PCB mixtures, the congener composition can be characterized using the mass fractions of all constituent congeners.
Maximum Likelihood Estimate	Estimate of a population parameter most likely to have produced the sample observations (U.S. EPA, 2015)
Reference mixture	Mixtures for which estimates of exposure levels associated with adverse effects (e.g., BMDs or ERCs from EPA's CatReg software), along with variance information for the estimates, can or have been derived.

Relative Toxicological Potency	Potency is a measure of chemical or toxicant activity expressed in terms of the amount required to produce an effect of given intensity. The relative toxicological potency of each PCB congener is an estimate of the congener's potency relative to the potency of an index congener, which is typically the congener with the highest known potency and is assigned a relative potency estimate of 1. For example, if exposure to a PCB congener causes a toxicological effect at a dose twice as large as the dose required for the index congener to cause the same level of the same effect, the relative potency of the congener would be 0.5 because it is half as potent as the index congener. For PCBs, relative toxicological potencies vary by health effect. For MiST, equal or unequal potencies can be assigned based on data availability.
Sufficient similarity	The case in which toxicological similarity between a candidate and reference mixture is judged sufficient to allow for use of the reference mixture's dose-response data in a risk assessment of the candidate mixture. Toxicological similarity exists between mixtures with components that are not very different in toxicological potency and are in about the same proportions such that the toxicological consequences of exposure to the two mixtures or their components are nearly identical (U.S. EPA, 2000).
Toxicological surrogate	A chemical or mixture with toxicological data sufficient for use in risk assessment that is used to support risk assessment of a related chemical or mixture for which data are limited or unavailable.

8.0 REFERENCES

Bull, RJ; Rice, G; Teuschler, LK. (2009). Determinants of whether or not mixtures of disinfection by-products are similar [Review]. Journal of Toxicology and Environmental Health, Part A: Current Issues 72: 437-460. http://dx.doi.org/10.1080/15287390802608916

Catlin, NR; Collins, BJ; Auerbach, SS; Ferguson, SS; Harnly, JM; Gennings, C, et al. (2018). How similar is similar enough? A sufficient similarity case study with Ginkgo biloba extract. Food and Chemical Toxicology 118: 328-339. http://dx.doi.org/10.1016/j.fct.2018.05.013

Feder, PI; Ma, ZJ; Bull, RJ; Teuschler, LK; Rice, G. (2009a). Evaluating sufficient similarity for drinking-water disinfection by-product (DBP) mixtures with bootstrap hypothesis test procedures [Review]. Journal of Toxicology and Environmental Health, Part A: Current Issues 72: 494-504. http://dx.doi.org/10.1080/15287390802608981

Feder, PI; Ma, ZJ; Bull, RJ; Teuschler, LK; Schenck, KM; Simmons, JE, et al. (2009b). Evaluating sufficient similarity for disinfection by-product (DBP) mixtures: Multivariate statistical procedures [Review]. Journal of Toxicology and Environmental Health, Part A: Current Issues 72: 468-481. http://dx.doi.org/10.1080/15287390802608965

Gennings, C. (2017). Expert assistance for the IRIS Draft Toxicological Review of PCBs - effects other than cancer: Testing of mixtures similarity tool. (EPA Contract No. EP-C-14-001). Research Triangle Park, NC: U.S. Environmental Protection Agency.

Holland, EB; Pessah, IN. (2021). Non-dioxin-like polychlorinated biphenyl neurotoxic equivalents found in environmental and human samples. Regulatory Toxicology and Pharmacology 120: 104842. http://dx.doi.org/https://doi.org/10.1016/j.yrtph.2020.104842

Hong, T; Avanasi, R; Henning, C. (2016). ICF Memorandum Re: PCB Similarity tool, assumption used to estimate standard error of the estimated distance, and tool testing. Research Triangle Park, NC: U.S. Environmental Protection Agency.

Mao, S; Zhang, G; Li, J; Geng, X; Wang, J; Zhao, S, et al. (2020). Occurrence and sources of PCBs, PCNs, and HCB in the atmosphere at a regional background site in east China: Implications for combustion sources: Supplementary material [Supplemental Data]. Environmental Pollution 262. https://doi.org/10.1016/j.envpol.2020.114267

Marshall, S; Gennings, C; Teuschler, L; Stork, L; Tornero-Vele, R; Crofton, K, et al. (2013). An empirical approach to sufficient similarity: Combining exposure data and mixtures toxicology data. Risk Analysis 33: 1582-1595. https://doi.org/10.1111/risa.12015

Murray, F; Roth, RN; Nicolich, MJ; Gray, TM; Simpson, BJ. (2013). The relationship between developmental toxicity and aromatic-ring class profile of high-boiling petroleum substances. Regulatory Toxicology and Pharmacology 67: S46-S59. http://dx.doi.org/10.1016/j.yrtph.2013.05.003

Prithiviraj, B; Chakraborty, P. (2020). Atmospheric polychlorinated biphenyls from an urban site near informal electronic waste recycling area and a suburban site of Chennai city, India. Science of the Total Environment 710. https://doi.org/10.1016/j.scitotenv.2019.135526

Rice, GE; Teuschler, LK; Bull, RJ; Simmons, JE; Feder, PI. (2009). Evaluating the similarity of complex drinking-water disinfection by-product mixtures: Overview of the issues [Review]. Journal of Toxicology and Environmental Health, Part A: Current Issues 72: 429-436. http://dx.doi.org/10.1080/15287390802608890

Ryan, KR; Huang, MC; Ferguson, SS; Waidyanatha, S; Ramaiahgari, S; Rice, JR, et al. (2019). Evaluating sufficient similarity of botanical dietary supplements: Combining chemical and in vitro biological data. Toxicological Sciences 172: 316-329. http://dx.doi.org/10.1093/toxsci/kfz189

Stork, L; Gennings, C; Carter, W; Teuschler, L; Carney, E. (2008). Empirical evaluation of sufficient similarity in dose—Response for environmental risk assessment of chemical mixtures. Journal of Agricultural, Biological and Environmental Statistics 13. https://doi.org/10.1198/108571108X336304

Trgovcich, J; Glen, G; Henning, C; Black, S; Silva, R. (2021). ICF report documenting MiST revisions and validation tests. ICF.

U.S. EPA (U.S. Environmental Protection Agency). (2000). Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures (EPA/630/R-00/002). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=20533

U.S. EPA (U.S. Environmental Protection Agency). (2010). Recommended toxicity equivalence factors (TEFs) for human health risk assessments of 2,3,7,8-tetrachlorodibenzo-p-dioxin and dioxin-like compounds [EPA Report]. (EPA/100/R-10/005). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/sites/production/files/2013-09/documents/tefs-for-dioxin-epa-00-r-10-005-final.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance. (EPA/100/R-12/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/risk/benchmark-dose-technical-guidance

U.S. EPA (U.S. Environmental Protection Agency). (2015). Benchmark Dose Software (BMDS) glossary of terms [Website]. https://sor.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&vocabName=BMDS%20Glossary&uid=1821254&taxonomyName=BMDS%20Glossary#formTop

U.S. EPA (U.S. Environmental Protection Agency). (2017). Categorical regression (CATREG) user guide, version 3.1.0.7. BMDS. https://www.epa.gov/sites/production/files/2016-03/documents/catreg_user_guide.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2019). Systematic Review Protocol for the Polychlorinated Biphenyls (PCBs) Noncancer IRIS Assessment Preliminary Assessment Materials. (EPA/635/R-19/201). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Integrated Risk Information System. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=237359

U.S. EPA (U.S. Environmental Protection Agency). (2020). Benchmark Dose Software (BMDS) Version 3.2 User Guide (EPA/600/R-20/216). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development. https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349749&Lab=CPHEA

Vorkamp, K. (2015). An overlooked environmental issue? A review of the inadvertent formation of PCB-11 and other PCB congeners and their occurrence in consumer products and in the environment [Review]. Science of the Total Environment 541: 1463-1476. http://dx.doi.org/10.1016/j.scitotenv.2015.10.019