DASS App User Guide

Table of Contents

Release Notes

20 October 2023

Version 1.1

- A demo data set has been added so that users may try using the app without the need to upload their own data.
- A supplemental module has been added which allows derivation of confusion matrices and performance metrics against user-supplied reference data.
- Columns in the results table can be hidden allowing users to view only their columns of interest.
- User does not need to reload app to upload new data.
- Collapsible panels, updated scroll behavior.

About

The DASS App is a user-friendly web tool for applying defined approaches on skin sensitization (DASS) that are described in Guideline No. 497 from the Organisation for Economic Co-operation and Development (OECD 2021) and the U.S. Environmental Protection Agency (EPA) Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing (EPA 2018). The DASS App predicts both skin sensitization hazard (either a sensitizer or non-sensitizer) and potency for chemicals of interest. For assigning potency predictions to query chemicals, the DASS App uses categories established by the United Nations Globally Harmonized System for Classification and Labelling of Chemicals (GHS).

The defined approaches (DAs) integrate data from in vitro assays that represent key events (KEs) in the Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins (OECD 2014) and in silico hazard predictions. Each KE is represented by a validated OECD test method:

1. The direct peptide reactivity assay (DPRA)(OECD 2022a) maps to KE 1, protein binding.
2. The KeratinoSens^{TM} assay (OECD 2022b) maps to KE 2, keratinocyte activation.
3. The human cell line activation test (h-CLAT) (OECD2022c) maps to KE 3, dendritic cell activation.

In addition, one of the DAs, the Integrated Testing Strategy (ITS) method, integrates an in silico prediction of skin sensitization potential generated by either Derek Nexus or the OECD QSAR Toolbox (Yordanova et al. 2019).

The app runs the DA analysis using data from a file you provide; data templates in Excel and text format are provided to help you correctly format your data. No data is provided by or retained by the app, and it can thus be used to analyze data from substances in development.

This app was developed by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). Version 1.1 was tested on Windows 10 using R v.4.2.2(R Core Team 2022) and Google Chrome v.111. The DASS App is available at https://ntp.niehs.nih.gov/go/952311.

For more information or to report a problem with the app, please contact NICEATM at ICE-support@niehs.nih.gov.

Step 1: Select Defined Approaches

In Step 1, select the DAs to be applied. Click the green information buttons next to the DA name to view a description of the DA and the test methods required to implement the DA.

2 out of 3 DA

The 2 out of 3 (2o3) DA predicts skin sensitization hazard based on KEs 1–3. The prediction is based on results from the DPRA, KeratinoSens, and h-CLAT, with at least two concordant results from the three assays determining the final prediction as a sensitizer or non-sensitizer (

Figure 1. Diagram of the 2o3 DA data interpretation procedure). If there are only results from two assays and the results are discordant, the chemical can't be classified, and the evaluation will return an "Inconclusive" result.

2o3 does not predict GHS potency category.

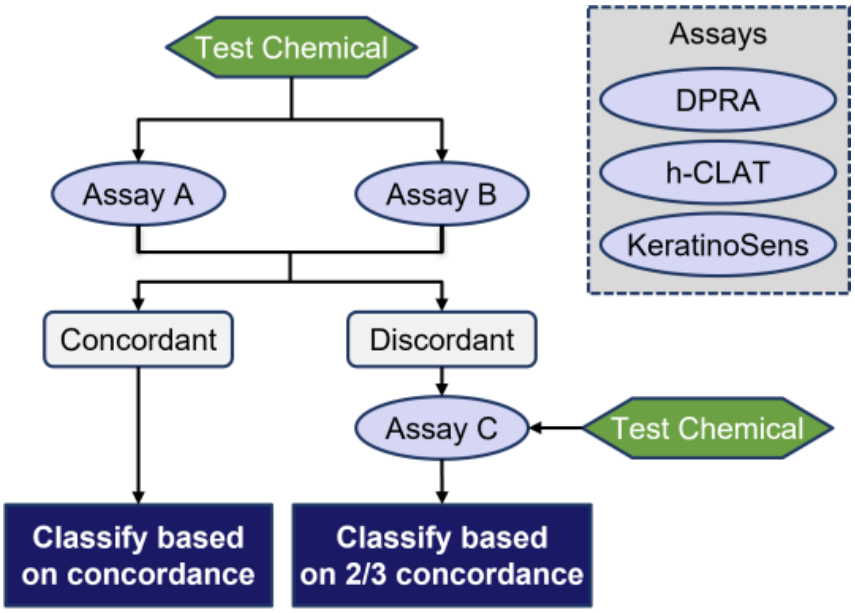

Figure 1. Diagram of the 2o3 DA data interpretation procedure

Integrated Testing Strategy DA

The ITS DA predicts skin sensitization hazard and GHS potency category based on KEs 1 and 3 and in silico predictions. Chemicals are scored using the minimum induction threshold (MIT) from the h-CLAT, the mean percent depletion for both cysteine and lysine in the DPRA, and hazard predictions from either Derek Nexus or the OECD QSAR Toolbox (Yordanova et al. 2019). The summed scores are used to predict chemical hazard and potency (**Error! Reference source not found.**).

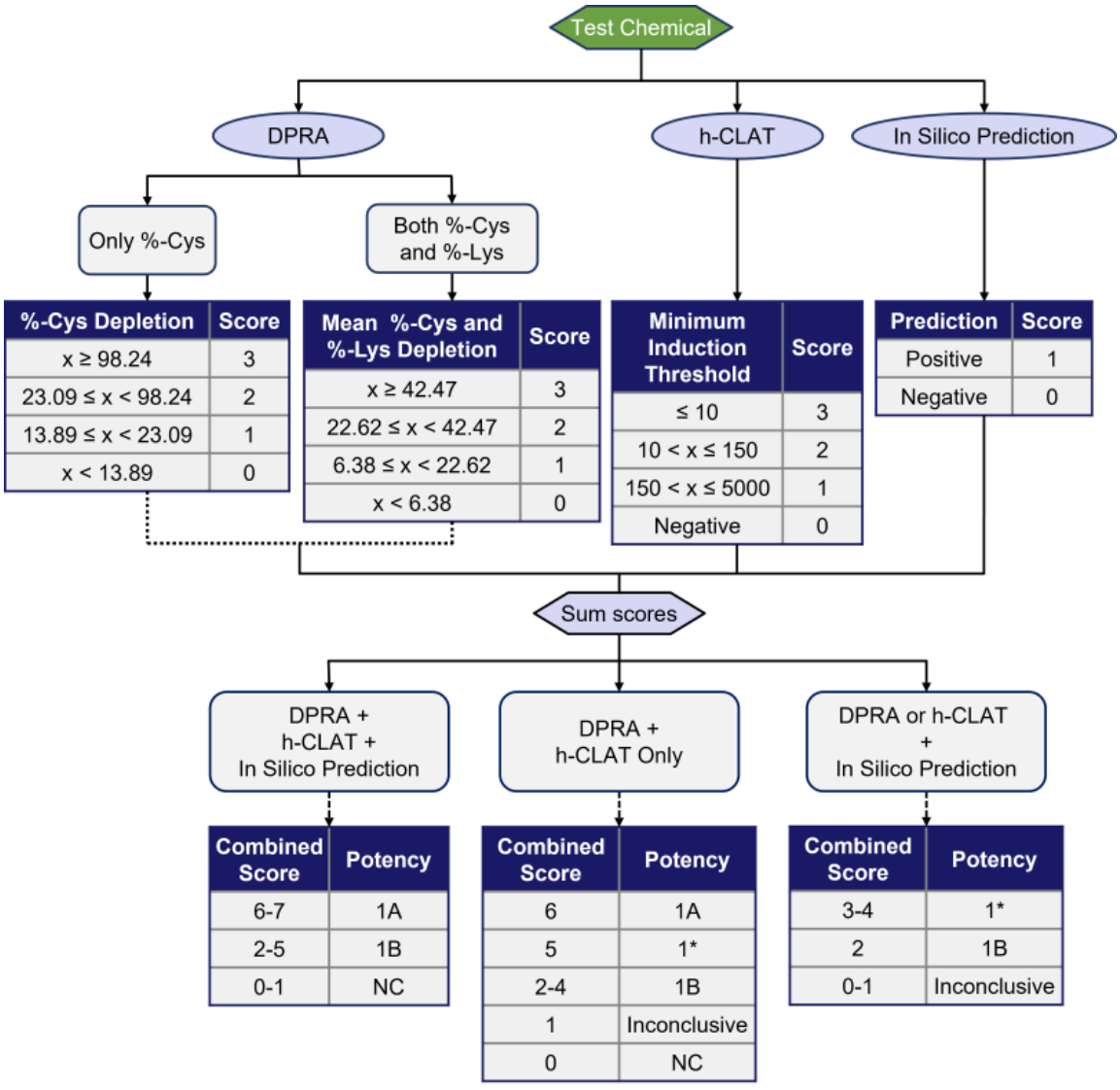

Figure 2. Diagram of ITS DA data interpretation procedure

Key Event 3/1 Sequential Testing Strategy DA

The KE 3/1 Sequential Testing Strategy (STS) predicts skin sensitization hazard and potency based on KEs 1 and 3. If the h-CLAT predicts that a chemical will be a sensitizer, then the hazard and potency categories are determined by the h-CLAT MIT. If the h-CLAT predicts the chemical to be a non-sensitizer, then the hazard and potency categories are determined by DPRA results (Figure 3.).

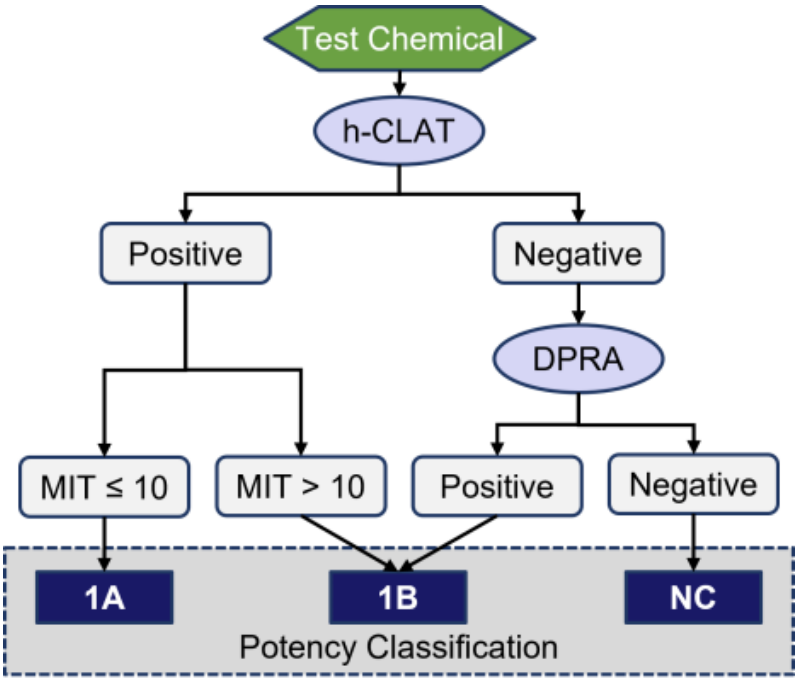

Figure 3. Diagram of KE 3/1 STS DA data interpretation procedure

Step 2: Upload Data

To upload the data, first click “Browse” and select your file using the dialog box that opens. Then click “Upload” to load the data into the app. Prior to uploading data, ensure that the data meet the data and formatting requirements.

The app provides a demo data set that can be used to build an understanding of how to use the app. Check the box labeled “Use demo data” to load the demo data set.

Data and Formatting Requirements

General

1. The file can be comma-delimited (.csv), tab-delimited (.tsv,.txt), or a Microsoft Excel workbook (.xls,.xlsx).
2. Data should be in a tabular format with each row corresponding to a single substance and a column for each required assay endpoint.
3. The first row should contain column names. Column names must be unique.
4. Missing values should be indicated by a blank cell or as 'NA' (without quotes).

Assay Endpoints

Each assay endpoint that is required for implementing the DAs should have a column that is formatted according to the formatting requirements shown in

Table 1. Values that do not meet the assay endpoint requirements will be treated as missing data and will not be used to derive predictions.

Table 1. Data format requirements for assay endpoints.

Assay	Endpoint	Format Requirements	2o3	ITS	STS
DPRA	Call	• Positive outcomes should be indicated by "sensitizer", "sensitiser", "active", "a", "positive", "pos", "p", or "1". • Negative outcomes should be indicated by "non-sensitizer", "non-sensitiser", "inactive", "i", "negative", "neg", "n", or "0".	X		X
DPRA	%-Cysteine Depletion	• Numeric values only. • No symbols.	O	X	O
DPRA	%-Lysine Depletion	• Numeric values only. • No symbols.	O	X	O
h-CLAT	Call	• Positive outcomes should be indicated by "sensitizer", "sensitiser", "active", "a", "positive", "pos", "p", or "1". • Negative outcomes should be indicated by "non-sensitizer", "non-sensitiser", "inactive", "i", "negative", "neg", "n", or "0".	X		
h-CLAT	MIT	• For positive h-CLAT outcomes, numeric values only. No symbols. • Indicate negative h-CLAT outcomes with "non-sensitizer", "non-sensitiser", "Inf", "i", "inactive", "n", "neg", or "negative".		X	X
KeratinoSens	Call	• Positive outcomes should be indicated by "sensitizer", "sensitiser", "active", "a", "positive", "pos", "p", or "1". • Negative outcomes should be indicated by "non-sensitizer", "non-sensitiser", "inactive", "i", "negative", "neg", "n", or "0".	X		
In Silico Prediction	Call	• Positive outcomes should be indicated by "sensitizer", "sensitiser", "active", "a", "positive", "pos", "p", or "1". • Negative outcomes should be indicated by "non-sensitizer", "non-sensitiser", "inactive", "i", "negative", "neg", "n", or "0".		X	
In Silico Prediction	Applicability Domain	• Predictions within the applicability domain should be indicated by "1" or "In". • Predictions outside the applicability domain should be indicated by "0" or "Out". These will be omitted from analysis.		X	

[Columns 4–6 indicate the DAs that require a given endpoint. Requirements are not case sensitive. X: the DA requires the endpoint. O: This endpoint can be used to derive a required call endpoint.]

Data Template

A downloadable data template is provided in tab-delimited and Excel format. The template is filled with example data that must be deleted before entering your data. The template contains a column for CASRNs and columns for every possible assay endpoint. Additional columns (e.g., other chemical identifiers) can be added to the file if needed. If an assay endpoint will not be used, the corresponding column can be deleted but that is not required.

Use of the template is not required.

Complete Data Upload

If you have uploaded an Excel file, select the appropriate worksheet from the dropdown menu. The data will be displayed for your review. Use the scroll bars on the right side and bottom of the display window to view your entire data set. Once you have reviewed your data, click “Continue” to proceed to the next step.

Step 3: Select Data Columns for Predictions

After clicking “Continue” in Step 2, the Step 3 section will expand, allowing you to select data to be used for the DA analysis. All assay endpoints required by the selected DAs will be shown. Dropdown lists under each assay endpoint contain all column names from the uploaded data. Use the dropdown lists to select the name of the column corresponding to each assay endpoint. Each column should only be selected once. The app will automatically select a column for an endpoint if that column name matches the corresponding column name in the data template.

Click the green information buttons next to the assay endpoint names to view information about the endpoints and column formatting requirements.

When you are done selecting columns, click “Done” to proceed to the next step. The app will not allow you to proceed unless you have specified a column from your uploaded data for each required input.

Step 4: Review Selection

After clicking “Done” in Step 3, the selected columns will be evaluated against the data and formatting requirements. A table with three columns will be displayed allowing you to review your inputs. The first column, “Variable”, contains the name of the assay endpoint. The second column, “Selected Column”, contains the name of the data column that was selected in the previous step. If the data column contains any invalid values, a warning message will appear at the top of the section, and the third column of the table, “Flag”, will contain text noting any requirements for the specific variable that your data have violated. A warning message will also appear at the top of the section if a unique column was selected for more than one assay endpoint.

Verify that the selected columns are correct. If needed, return to Step 3 to change column selections. Clicking “Done” in Step 3 will update the table.

Rows corresponding to a flagged user column will contain orange text. If these appear, first review your selection in Step 3 to verify whether the flag was triggered by an incorrect column selection. If the flag was triggered by incorrect formatting in the source data:

1. Open your Excel or text file providing the source data and review the formatting of the column corresponding to the flagged data.
2. Make the needed corrections to the data in the source file.
3. Return to DASS App Step 2, upload your corrected data file, and click “Continue”.

After reviewing the warnings and flags, click “Run” to generate the DASS predictions.

The DASS App will allow you to proceed with generating predictions if your data have unresolved flags. Any invalid values will be marked as missing data and will not be used for predictions.

Step 5: Results

After the DAs have been applied, a table with the DASS predictions will be displayed in the Results panel. Use the scroll bars on the right side and bottom of the display window to view the entire results table. To hide or show specific columns, click on “Column Visibility” and use the menu to select or deselect columns. You can also use the Download Results dropdown menu to download the results table as a tab-delimited or Excel file, which may allow easier viewing of the entire table.

- The data columns that were selected in Step 3 are highlighted in yellow. The column names are annotated with an asterisk. These data columns are reformatted for use in the DAs.
- The reformatted columns are appended to the original data columns and highlighted in pink. The corresponding column names begin with “Input”. For values that were calculated by the app, the column name will also end with “Calculated”.
- The DASS predictions are appended to the uploaded data and are highlighted in blue. If the ITS DA was selected, then the individual ITS scores are also appended and highlighted in blue. The corresponding column names begin with ”DA” and the abbreviated name of the DA. The DASS prediction columns are described in Table 2.

Table 2. Descriptions of prediction columns in the results table.

Column Name	Description
DA ITS h-CLAT Score	ITS score for the h-CLAT MIT outcomes.
DA ITS DPRA Score	ITS score for the DPRA %-cysteine and %-lysine outcomes.
DA ITS in Silico Score	ITS score for the in silico outcome.
DA ITS Total Score	Final score from the ITS DA used to predict call and potency.
DA ITS Call	Hazard prediction from the ITS DA.
DA ITS Potency	Potency prediction from the ITS DA.
DA 2o3 Call	Hazard prediction from the 2o3 DA.
DA KE 3/1 STS Call	Hazard prediction from the KE 3/1 STS DA.

DA KE 3/1 STS Potency	Potency prediction from the KE 3/1 STS DA.

Error! Not a valid bookmark self-reference.3 contains a description of the input columns and how they were reformatted. For example, if you selected only the 2o3 DA, then the column "h-CLAT Call Input" would be added to the results, but the column "h-CLAT MIT Input" would not be added because it is not a required endpoint for the 2o3 DA.

Table 3. Descriptions of columns in the results that contain reformatted user column data.

Column Name	Description
DPRA Call Input	Positive outcomes are set to 1 and negative outcomes are set to 0. Any invalid values are set to NA.
DPRA %-C Depletion Input	Only numeric values are retained. Any invalid values are set to NA.
DPRA %-K Depletion Input	Only numeric values are retained. Any invalid values are set to NA.
DPRA Mean (Calculated)	The calculated mean of DPRA %-C and %-K depletion values, where any negative values are set to 0 when calculating the mean. DPRA mean is used when the ITS DA is selected or when DPRA hazard is derived from depletion values.
DPRA Call Input (Calculated)	The outcome derived from DPRA Mean (Calculated) when 2o3 or KE 3/1 STS DAs are selected, and the user provides DPRA %-C and %-K depletion columns instead of a DPRA hazard column.
h-CLAT Call Input	Positive outcomes are set to 1 and negative outcomes are set to 0. Any invalid values are set to NA.
h-CLAT MIT Input	Numeric values for positive h-CLAT outcomes are retained. In addition, negative h-CLAT outcomes are set to Inf. Any invalid values are set to NA.
Keratinosens(TM) Call Input	Positive outcomes are set to 1 and negative outcomes are set to 0. Any invalid values are set to NA.
In Silico Call Input	Positive outcomes are set to 1 and negative outcomes are set to 0. Any invalid values are set to NA.
In Silico Applicability Domain Input	Predictions in the applicability domain are set to 1 and predictions outside the applicability domain are set to 0. Any invalid values are set to NA.

Supplemental: Compare Results

After the DAs have been applied, the Supplemental panel will display options for evaluating the results against reference data. Results can be evaluated for either hazard or potency predictions.

The dropdown list for prediction columns contains the column names from the results table that correspond to the DA predictions. Use the dropdown list to select the DA results you want to evaluate. The dropdown list for reference columns contains all column names from the uploaded data. Use the dropdown list to select the columns in your data that contain reference data. The app does not provide reference data. Reference data must be included in your

original data upload. Formatting requirements for the reference columns are provided in Table 4.

Table 4. Data format requirements for reference values.

Type of Comparison	Format Requirements
Hazard	• Positive outcomes should be indicated by “sensitizer”, “sensitiser”, “active”, “a”, “positive”, “pos”, “p”, or “1”. • Negative outcomes should be indicated by “non-sensitizer”, “non-sensitiser”, “inactive”, “i”, “negative”, “neg”, “n”, or “0”. • Missing values should be blank or labeled as “NA”.
Potency	• Potency should be one of 3 values corresponding to GHS potency categories: 1A, 1B, or NC • Missing values should be blank or labeled as “NA”.

When you are done selecting columns, click “Compare” to generate results. Output is produced for every pairwise comparison between the selected prediction and reference columns. Output is uniquely labeled by the names of the selected columns for a given comparison. Use the “Select Output” dropdown list to select the output you want to view.

For each pairwise comparison, a confusion matrix and performance metric table are shown. Click “Download” to open a menu with options for exporting the results. Confusion matrices and performance metrics can be downloaded as a PDF. The results can be downloaded in tabular format as an Excel file or text file.

Performance Metric Derivation

Hazard

- N = The number of valid reference values
- Accuracy = (True positives + True negatives) / (All positives + All negatives)
- Balanced Accuracy = (True positive rate + True negative rate)/2
- F1 Score = (2×True positives) / (2×True positives + False positives + False negatives)
- True Positive Rate (Sensitivity) = True positives / All positives
- False Positive Rate = False positives / All positives
- True Negative Rate (Specificity) = True negatives / All negatives
- False Negative Rate = False negatives / All negatives

Potency

- N = The number of valid reference values
- Accuracy = The percentage of predicted values equal to reference values

- Overpredicted = The percentage of predicted values with a more potent GHS category than the corresponding reference value
- Underpredicted = The percentage of predicted values with a less potent GHS category than the corresponding reference value

References

EPA. 2018. Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing Draft for Public Comment. EPA's Office of Chemical Safety and Pollution Prevention.

OECD. 2014. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. OECD Series on Testing and Assessment, No. 168. https://doi.org/10.1787/9789264221444-en.

OECD. 2021. Guideline No. 497: Defined Approaches on Skin Sensitisation. OECD Guidelines for the Testing of Chemicals, Section 4. https://doi.org/10.1787/b92879a4-en.

OECD. 2022a. Test No. 442C: In Chemico Skin Sensitisation Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins, OECD Guidelines for the Testing of Chemicals, Section 4. https://doi.org/10.1787/9789264229709-en.

OECD. 2022b. Test No. 442D: In Vitro Skin Sensitisation. ARE-Nrf2 Luciferase Test Method, OECD Guidelines for the Testing of Chemicals, Section 4. https://doi.org/10.1787/9789264229822-en.

OECD. 2022c. Test No. 442E: In Vitro Skin Sensitisation. In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation, OECD Guidelines for the Testing of Chemicals, Section 4. https://doi.org/10.1787/9789264264359-en.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing:Vienna, Austria.

Yordanova D, Schultz TW, Kuseva C, Tankova K, Ivanova H, Dermen I, et al. 2019. Automated and standardized workflows in the OECD QSAR Toolbox. Comput Toxicol 10:89–104. doi:https://doi.org/10.1016/j.comtox.2019.01.006.