Purposes and Uses of EPA's Toxicity ForeCaster (ToxCast™)

Q&A Document for ACC members prepared by the ACC LRI

Introduction

This Q&A document discusses and illustrates the purpose and uses of EPA's ToxCast™ data. It is intended to provide a technical summary of how to find, use, and interpret ToxCast data to facilitate ACC members' understanding of advanced approaches for biological profiling of chemicals, and the use of this information within chemical safety assessments. It is not intended to be a comprehensive guide to ToxCast, rather an introduction and resource of where to find more information. References and hyperlinks are provided for those interested in accessing the relevant literature and resources. It is envisioned that this document will be evergreen and will be updated on a periodic basis as relevant information on ToxCast is developed, as improved guidance for ToxCast becomes available, and as new questions arise.

All screenshots and links are updated as of August 2022.

Contents

Q1. What is ToxCast?

ToxCast and Tox21 are high-throughput chemical screening programs designed in response to the 2007 NRC report "Toxicity Testing in the 21st Century: A Vision and a Strategy," which called for the use of *in vitro* assays to determine the effects of thousands of largely untested chemicals on "toxicity pathways." The ToxCast (Toxicity ForeCaster) program is run by the Center for Computational Toxicology and Exposure (CCTE, formerly the National Center for Computational Toxicology) within the EPA. According to a subsequent publication by the EPA, *In vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project.* The overall goals of this program are to identify *in vitro* assays and responses that are relevant to *in vivo* toxicity, to develop predictive models built on the results of multiple assays and chemical properties, and to use these assays and models to screen environmental chemicals that have little or no available toxicity data and prioritize them for further testing. The term Tox21 is often used to describe the concept of toxicity testing in the 21st century, but it is also the name of the formal collaboration established to address this issue of developing new toxicity testing strategies. Tox21 leverages the expertise of several federal agencies: US EPA (ToxCast), National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS, formerly the NCGC), National Institute of Environmental Health Sciences (NIEHS) National Toxicology Program (NTP), and the Food and Drug Administration (FDA).

EPA has invested time into developing additional trainings on NAMs including the CompTox Chemicals Dashboard which now houses the ToxCast Data. To view these trainings visit https://www.epa.gov/chemical-research/new-approach-methods-nams-training.

Within the ToxCast program, chemical testing is coordinated by the CCTE, with data being collected by outside vendors and then analyzed and shared by the CCTE. As of the completion of Phase III in 2018, ToxCast has screened over 4,500 chemicals in more than 700 high-throughput assays that cover many high-level cell responses and more than 300 signaling pathways.

Chemical screening is also carried out by NCATS, collecting data for over 8,000 chemicals in more than 50 assays. This data is analyzed by the NTP, and data is available in the Tox21 Toolbox. This raw data was also provided to EPA's CCTE to be analyzed through their data processing pipeline, so it is included in the CompTox Dashboard on the "TOXCAST/TOX21" subtab.

ToxCast is an ongoing project, therefore versioning is quite important. The testing can be separated into phases and the phase indicates which chemicals and assays were tested. Chemicals and assays can be added or removed, and the data analysis methods are also evolving and being improved. Updating the data derived from a particular set of chemicals and assays results in a new "Data Release." Phases I-III have been completed with the first data release in 2010 and the most recent version being the invitroDBv3.3 database released in March 2020. It is recommended to use the most up-to-date data available, and always be mindful of the data version used for publications when making comparisons.

Q2. What can I do with ToxCast Data?

ToxCast data can be used to better insight into the bioactivity of chemicals and the potential mechanistic pathways that they act on, with caveats as described in Question 8. While ToxCast data provide information on bioactivity, such data do not, on their own, indicate hazard or an adverse effect *in vivo*. ToxCast assay results may be useful in tiered testing and prioritization (see, for example, *Incorporating New Technologies Into Toxicity Testing and Risk Assessment: Moving From 21st Century Vision to a Data-Driven Framework*, and *Developing context appropriate toxicity testing approaches using new alternative methods (NAMs).* ToxCast data along with other knowledge (e.g., genomics, structural activity relationships, etc.) can help accelerate Adverse Outcome Pathways (AOP) discovery, development, and evaluation.

In vitro to *in vivo* extrapolation (IVIVE), pioneered by the ACC LRI program, has been used to convert *in vitro* bioactivity concentrations in ToxCast assays into external equivalent doses, which can then be compared to predicted or measured human exposures to enable risk-based decision making based on margins of exposures. *Incorporating High-Throughput Exposure Predictions with Dosimetry-Adjusted In vitro Bioactivity to Inform Chemical Toxicity Testing* illustrates this approach.

For certain receptor mediated responses, ToxCast results have been shown to be useful in predicting specific *in vivo* responses. For example, Brown and colleagues show in *Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model* that a ToxCast-derived estrogen receptor (ER) bioactivity model can be used as a measure of relative bioactivity, and ToxCast assay results applied to this model can be used as an alternative to the *in vivo* uterotrophic assay. (Note: the ER model has been subsequently refined; see *On Selecting a Minimal Set of In vitro Assays to Reliably Determine Estrogen Agonist Activity.)* Kleinstreuer and colleagues have also published *Development and Validation of a Computational Model for Androgen Receptor Activity* based on ToxCast assays. The strengths and limitations of IVIVE methods have been further explored in *In vitro to in vivo Extrapolation for High Throughput Prioritization and Decision Making.* In the publication *An Exposure:activity Profiling Method for Interpreting High-Throughput Screening Data for Estrogenic Activity--Proof of Concept*, the authors indicate how ER bioactivity can be evaluated using IVIVE and developing margins of exposure. In addition, in this same paper, the bioactivity to exposure ratio for the natural phytoestrogen is used to place these results into context. This method has subsequently been expanded in *Employing Dietary Comparators to Perform Risk Assessments for Anti-Androgens Without Using Animal Data* for substances with anti-androgenic activity.

ToxCast assays, however, have not performed well in predicting whether or not a chemical could be classified as a carcinogen. In the study *How Well Can Carcinogenicity Be Predicted by High Throughput "Characteristics of Carcinogens" Mechanistic Data*? Becker and co-workers showed that ToxCast results for the so-called key characteristics of carcinogens put forward by IARC were no better than chance in predicting cancer hazard. That said, subsequent research presented in *Utility of In vitro Bioactivity as a Lower Bound Estimate of In vivo Adverse Effect Levels and in Risk-Based Prioritization* has shown that while such data cannot reliably predict a specific *in vivo* hazard, using an *in vitro* point of departure (POD) derived from ToxCast results using IVIVE can be a health-protective POD for use in screening-level safety evaluations. *The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency* further discusses such POD approaches and applications.

Q3. What types of information does ToxCast data provide?

ToxCast provides a broad spectrum of in vitro bioactivity results on molecular targets or cellular pathways for the chemicals that it has analyzed. The information discussed below can be accessed primarily through EPA's CompTox Dashboard. Use the screenshots and information below to help guide you through the CompTox Dashboard.

Chemical information

Extensive information regarding chemical selection, procurement, management, and applicability domain is available in a recent publication and more background information about the chemical lists and procurement workflow can be found here: https://www.epa.gov/chemical-research/toxcast-chemicals. A full list of the ToxCast chemical screening library can be found here: https://comptox.epa.gov/dashboard/chemical_lists/TOXCAST.

The primary identifier for chemicals in the dashboard is the DTXSID, a unique chemical identifier created by CCTE. Chemicals can also be identified by their CAS number and chemical name.

Each chemical has been tested in a set of assays, and the resulting data has been processed through a data clean up pipeline and can be accessed via EPA's CompTox Dashboard. Within the dashboard, the ToxCast data is accessible for a given chemical within that chemical's Bioactivity tab. See Q4 for more details for accessing this data.

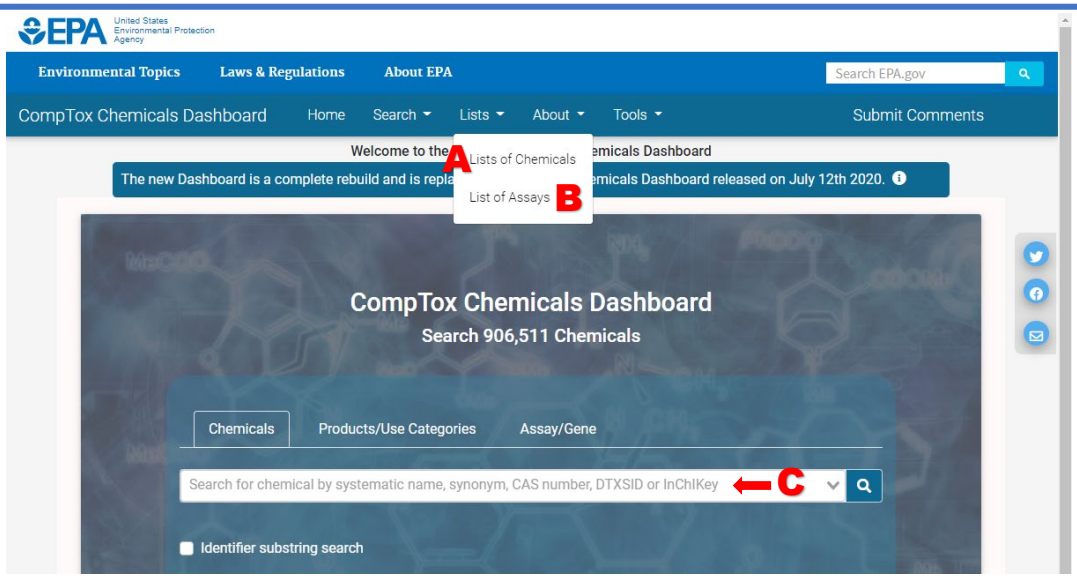
Assay information

Although EPA's CCTE coordinates testing, provides the chemical samples, and processes the data, independent laboratories or vendors mostly run the assays. The assay technologies are diverse and cover a wide variety of endpoints. Background information about the assay technologies, platform sources, protocols, and quality statistics can be found here: https://www.epa.gov/chemical-research/toxcast-data-generation-toxcast-assays

Assays are listed in the Dashboard by an abbreviated but informative annotation. The annotations are intended to provide: 1) assay identification information; 2) assay design information; 3) target information and 4) analysis information. For example, "CEETOX_H295R_ESTRADIOL_up" is from the vendor CEETOX, the assay is run in H295R (adrenocortical) cells, the hormone being measure is estradiol and the data was analyzed for activity in the "up" direction (looking to detect an increase in hormone level). Further explanations can be found in the following document from the data download page: Assay Annotation User Guide.

The document linked within the Details column in the Assay List after clicking "List of Assays" (See Exhibit 2) provides a variety of information about the selected assay. Annotations, parameters and descriptions about the assay technology and conditions are listed. Links to PubMed for any related publications are given. The data processing methods that were applied in the different levels of analysis for the particular assay are listed and described. A table shows the reagents and conditions used both for the cell culture (if applicable) and the assay for bioactivity. There are links to any Adverse Outcome Pathways (AOPs) that may exist for the intended target of the assay along with Assay Tags you can click to use as filters for the assay list. Graphs are used to present the summary statistics for all chemicals that were active in the chosen assay (binned by AC50 or scaled top value) and the numerical data is also provided. The AC50 is defined as the *in vitro* concentration at 50% of maximum activity. In addition to being found in the Assay List, you can access the information about an assay used for a specific chemical within the ToxCast: Summary subtab. After clicking the ToxCast: Summary subtab from the specific chemical's dashboard, click on the document icon under the "Modal" Column and you will be able to access the same assay information as described previously.

Exhibit 1: CompTox Dashboard Home Page



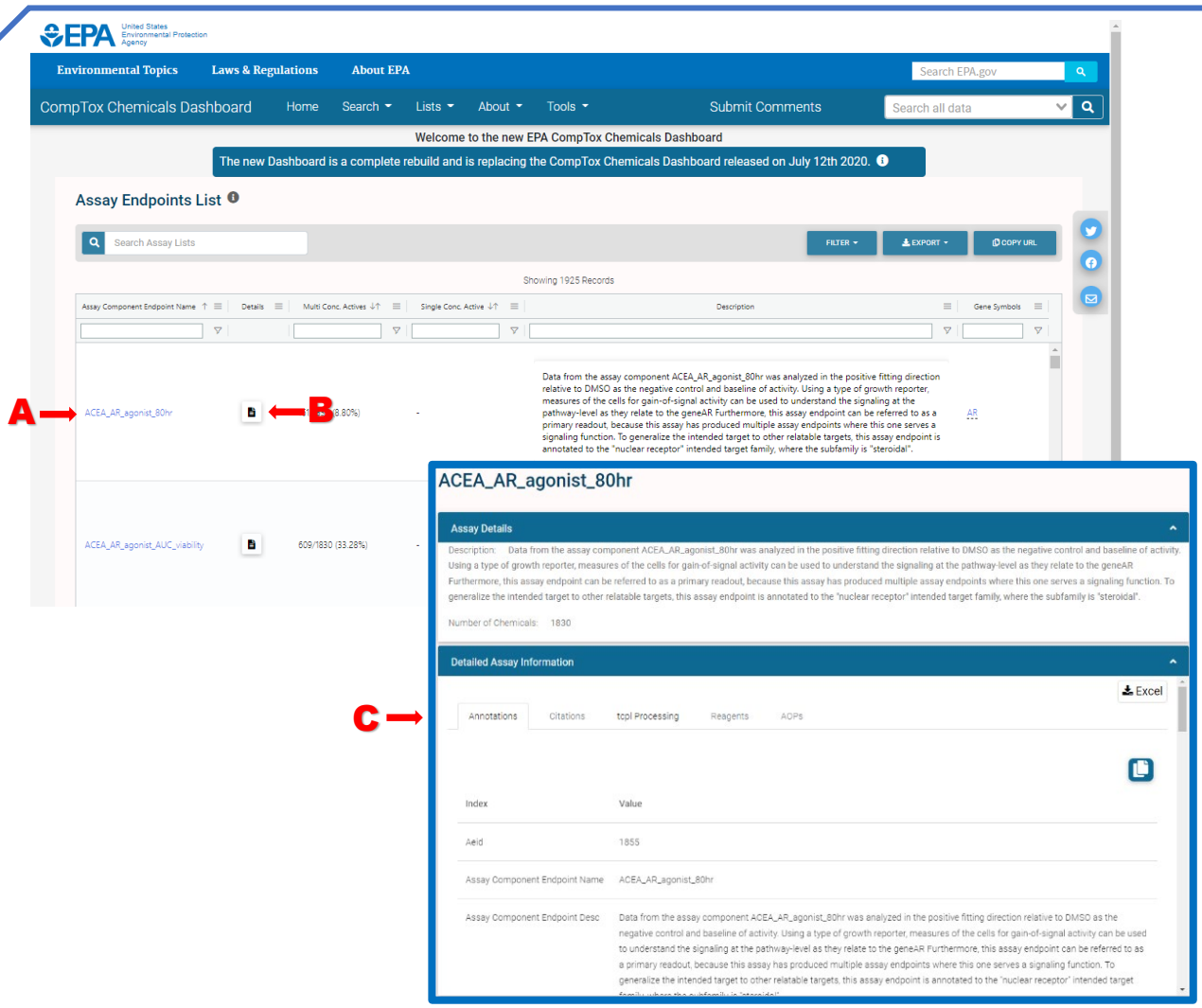
A – Click "List of Chemicals" to find a comprehensive list of chemicals included in the CompTox Dashboard.

B – Click "List of Assays" to find a comprehensive list of assays included in the CompTox Dashboard.

C – Use the search bar to search for a specific chemical quickly.

Source: https://comptox.epa.gov/dashboard

Exhibit 2: List of Assays Page



A – Clicking on the Assay name will open a search page with all chemicals that were analyzed using that assay.

B – Clicking on the document icon will open the details of that assay.

C – Additional details of the selected assay are shown by clicking on the Assay name.

Source: https://comptox.epa.gov/dashboard/assay_endpoints/

Q4. How can I access the ToxCast data?

All the curated data generated by the ToxCast and Tox21 programs is made freely available to the public. The data can be accessed in three primary ways, depending on the needs and abilities of the user: the interactive online CompTox Dashboard, downloadable summary files, or the NICEATM ICE 3.5 Dashboard. The most user-friendly way to access the data is directly through the CompTox Dashboard which allows for quick, easy access to data on single chemicals or assays.

The NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) Integrated Chemical Environment (ICE) is an online computational toxicology tool that incorporates curated data from multiple sources to provide resources to facilitate the safety assessment of chemicals. The original data used in ICE is the same raw data as

that of ToxCast but processed using a different analysis pipeline. For more information on ICE, please review the about page on their website, or in Question 5 below.

Additionally, for the most advanced or specialized users, the full database can be accessed through MySQL and the use of R programming scripts. This allows the user access to all levels of data (not just the final processed results) as well as the ability to manipulate the data and even the processing methods. Instructions for downloading and using the MySQL database and the R package to manipulate the data are available from the data download page in the ToxCast Data Pipeline Overview document and will not be discussed further here.

Using the CompTox Dashboard to Access ToxCast Data

Data in the Dashboard can be viewed in a few different ways:

- Find a chemical or assay of interest by searching for a Chemical (e.g., using CASRN number or name), an Assay Endpoint Name, or a Gene Symbol in the search bar. (See Exhibit 1).
- Find a chemical within the ToxCast_V3 list after selecting List of Chemicals. (See Exhibit 1).
- Find an assay after clicking on the List of Assays and review the list of chemicals that have been tested with that assay. (See Exhibit 1 and Exhibit 2).
    - Note: Not all the chemicals that populate will necessarily be from the ToxCast data set, as other data sets are included in the CompTox Dashboard.
- Accessing one of the various Chemical lists flagged as ToxCast within this list on the CompTox Dashboard.

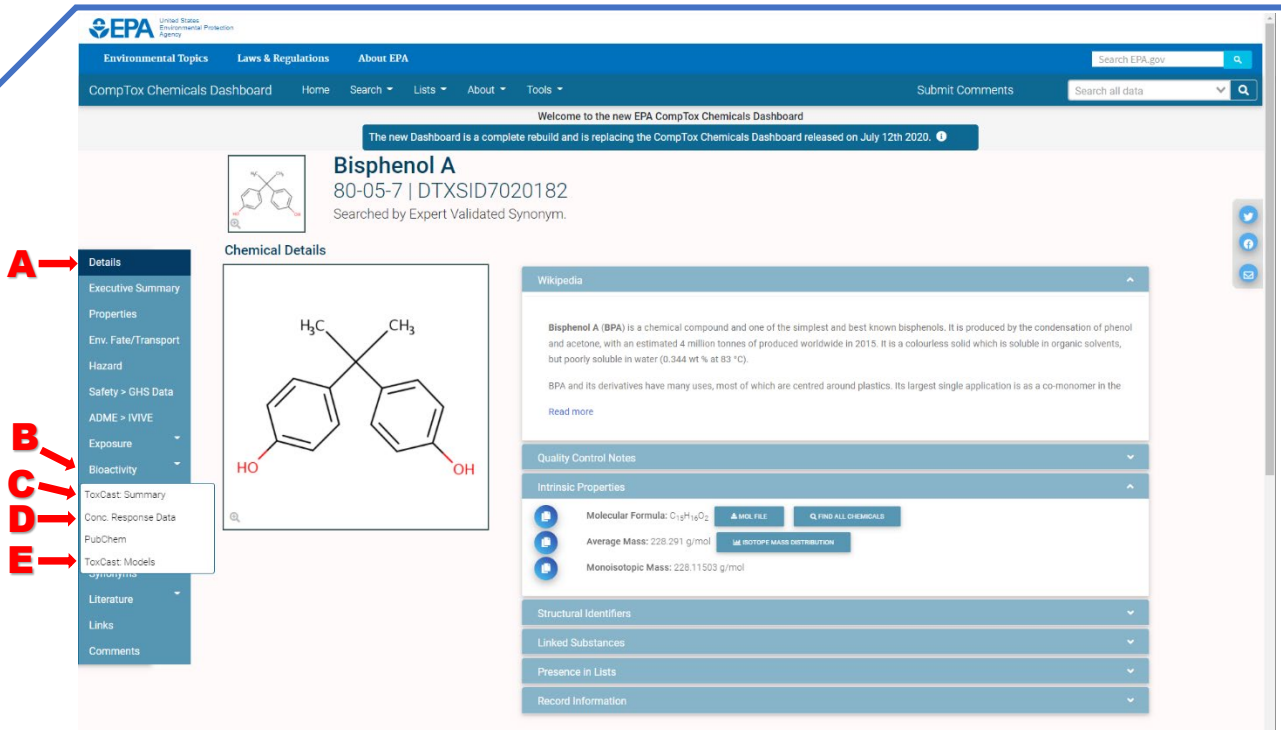
Once a chemical has been selected (by clicking on it), the Details tab for that chemical is automatically displayed. To find the ToxCast data, click on the Bioactivity tab (see Exhibit 3).

When available, hover over the  icon for more details on a given component of the dashboard. These are located throughout the Dashboard and contain useful information.

Chemical Details Tab

The Details tab is the landing point that appears after a chemical is selected, which provides several pieces of information including background, information on properties, structural identifiers, and information as to which lists the chemical appears on.

Exhibit 3: Example Chemical Information Page – BPA



A – The Details tab provides extensive information on the background of the selected chemical.

B – Bioactivity Tab Set – Houses ToxCast Data

C – ToxCast Summary Data tab shows all ToxCast data for the selected chemical.

D – Conc. Response tab allows you to review all assays for the selected chemical.

E – ToxCast Models tab

Source: https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID7020182

Bioactivity Tab Set

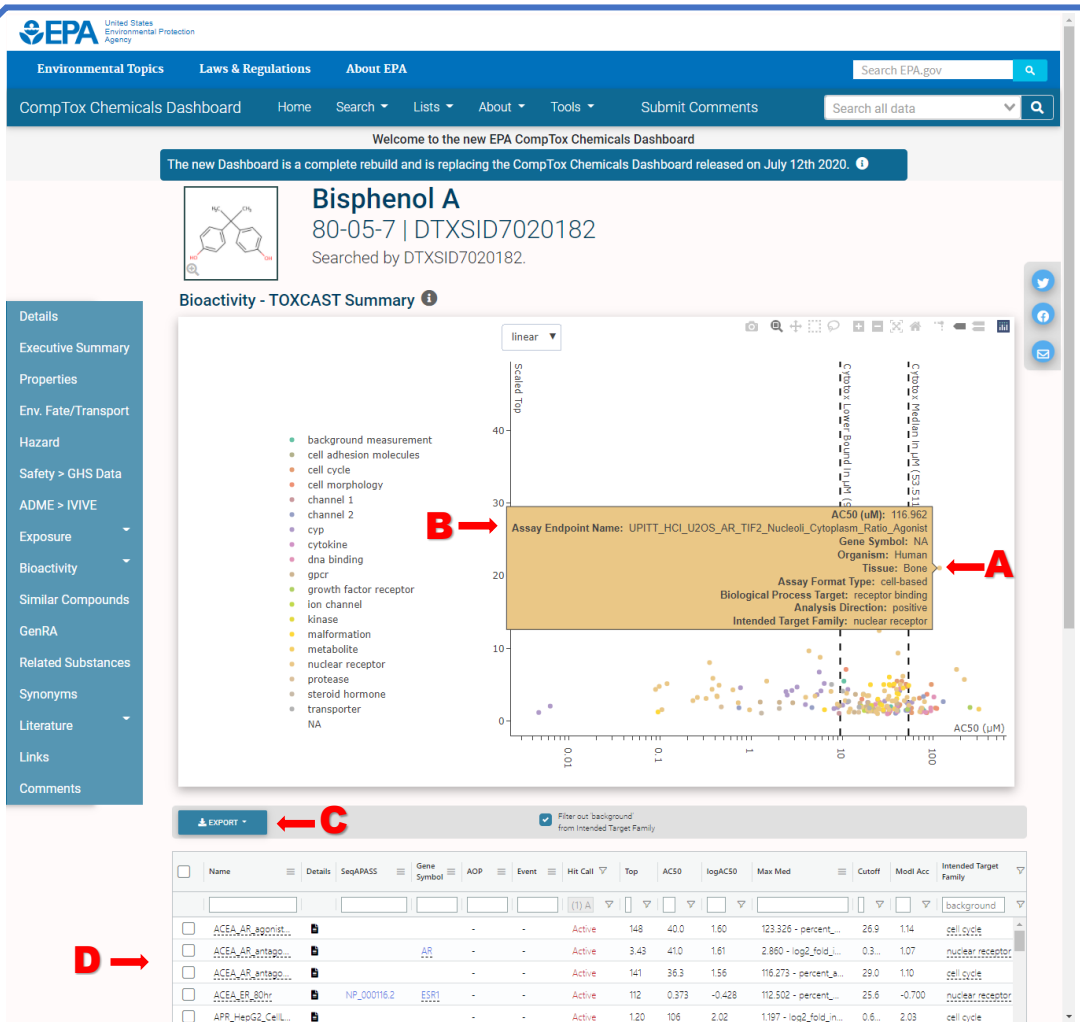
When the Bioactivity tab set is selected, the ToxCast: Summary subtab is automatically opened. The Bioactivity subtabs include ToxCast, Concentration Response, and PubChem datasets. The sections below describe what data each of the subtabs contain.

ToxCast: Summary Subtab

This subtab contains an overview of the AC50 data across assays for the selected chemical, as shown in Exhibit 4. The assays are grouped on the left based on the mechanism they are intended to explore and selecting or deselecting a grouping on the left will filter the graph accordingly. Individual points can be selected on the graph which will populate the box to the right of the screen with information about that assay. Below the graph, the assay data is also contained in a table. Once the desired endpoints or assays have been selected, both the data table and the image of the graph itself are able to be downloaded using the various download buttons within the tab.

The tips on this tab contain helpful information about what can be found on this tab.

Exhibit 4: ToxCast: Summary Subtab



A – Selected point (any of the points can be selected)

B – Selecting a specific point will pull up all relevant assay information here.

C – Data from this tab can be downloaded by clicking the EXPORT button.

D – Assay data for this chemical are listed below.

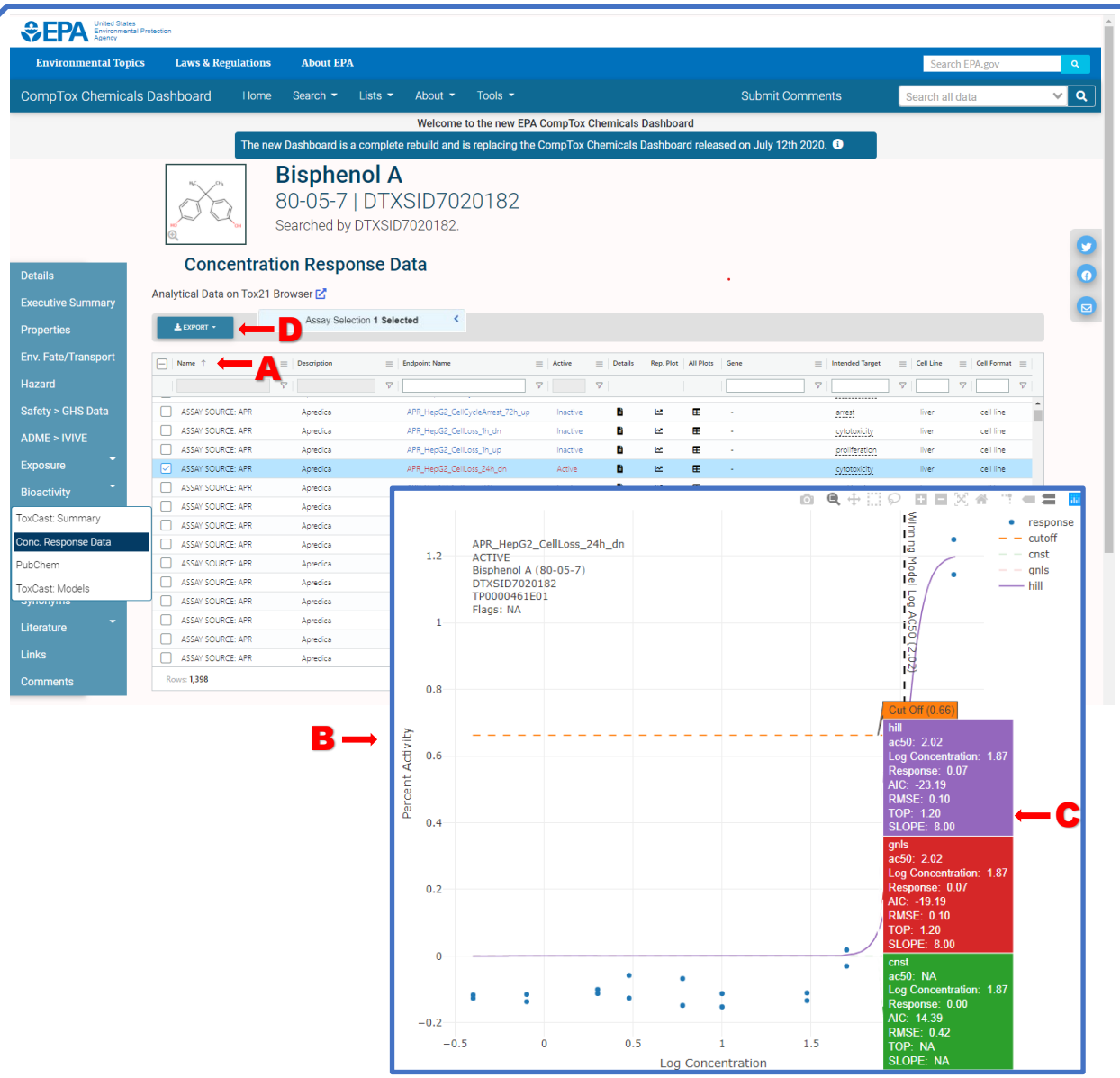
Source: https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID7020182

Conc. Response Data Subtab

The Conc. Response Data subtab provides the bulk of the data for each individual assay. As shown in Exhibit 5, the list of applicable assays that the chemical has been tested in, organized by vendor, populates in the column to the left. Selecting any one of those vendors will open a drop-down list of each individual assay that has been performed. Red assays are Active, and Blue Assays are inactive (see Question 5 for further explanation). Selecting an assay causes the bioactivity curve to populate to the right. Multiple assays can be selected at once. Below the curve a table also populates that contains the three models, which model won, and data about the curve including the AC50 value.

Note: Next to each assay in the list is a symbol. Hovering over this gives a brief summary and clicking on this provides more extensive valuable information about that assay.

Exhibit 5: Conc. Response Data Subtab



A – All assays associated with the selected chemical can be reviewed here and filtered by active or inactive status.

B – This is an example of a Dose Response curve that appears after clicking the plot icon for an assay.

C – Hover to show data and compare models.

D – The data selected can be downloaded as an Excel or .csv file.

Source: https://comptox.epa.gov/dashboard/chemical/concentration-response-data/DTXSID7020182

ToxCast Models Subtab

This subtab shows the chemical-specific results of predictive models based on the ToxCast data. Fully exploring this subtab is out of scope of this resource but hovering over the icon will give information about the predictive model including reference. The default for the "Display Bioactivity" plots associated with an assay is "Representative Samples Only," which produces only one chart per assay. To see all charts for an assay, one must deselect "Representative Samples Only."

Accessing the data through direct download

The ToxCast Data are available for download through the CompTox Dashboard and can be accessed through R scripts. Some examples of data that can be downloaded are as follows:

- The data for an individual chemical can be downloaded from the ToxCast Summary or the Conc. Response Data subtabs for that chemical.
- Data for a particular assay can be downloaded after selecting it from the List of Assays (See Exhibit 1 and Exhibit 2).

For more information on how to download ToxCast data and EPA's TCPL Scenarios in R, please review this helpful guide prepared by the EPA.

Q5. How is the ToxCast data curated?

The ToxCast Data Analysis Pipeline

The chemical screening data from Conc. Response are processed and modeled using scripts written in the R programming language. The ToxCast program developed the tcpl R package to process, normalize, model, qualify, flag, inspect and visualize the data. In depth information of the data analysis pipeline can be found on EPA's website, so only a brief overview will be provided here.

Raw data provided by a vendor or laboratory is processed, indexed, transformed, and normalized using standardized methods (described as Lvl 0 – 3 in the R package). Lvl 4 attempts to fit three models to the concentration-response data: constant, Hill, and gain-loss. The data may fit one or more of the models, or none at all. In Lvl 5, if any models were sufficiently fit, this chemical-assay pair is considered "active" (hit call = active). Then the winning model is chosen based on the lowest AIC (Akaike information criteria) value, which is used to indicate the model that best fits the data. The response cutoff, AC50 (concentration where half maximal activity occurs) and other parameters are calculated based on the winning model. For chemical-assay pairs where no models were sufficiently fit, the hit call is "inactive," and a winning model is not chosen. The final step is to assign "flags" or warnings to the data when methods in the data processing pipeline have identified possible false positive or false negative findings. Flags are discussed in more detail in Question 6.

In addition to the hit call, AC50, and other numerical parameters resulting from the data analysis, dose- response curves are available to visualize the data select the Conc. Response Data subtab under the Bioactivity tab (Exhibit 5) after selecting the assay(s) of choice. The dose-response curve will appear in the open panel with the AC50 indicated by a blue line on the graph and in the table below. All three models are shown in the table below, and the winning model is indicated as well.

The NICEATM ICE Database

As stated previously, the NICEATM ICE 3.5 Database incorporates the Tox21 data through the cHTS data set which contains data from assays run by laboratories participating in the Tox21 Consortium. The main difference between the data present in the CompTox Dashboard and ICE is that NICEATM used their own method of curating the data and evaluating quality including annotations based on their mechanistic target to simplify and to allow linkage to other molecular and cellular response pathways.

For more information, visit the NICEATM ICE About page, their site outlining the data curation process for the cHTS data, or this 2020 publication: *An integrated chemical environment with tools for chemical safety testing.*

Q6. How do I interpret ToxCast?

Interpretation of ToxCast Results for a Single Chemical

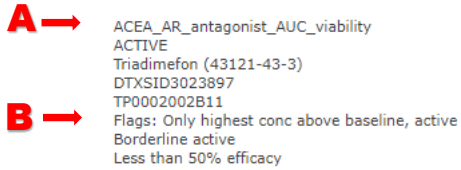
ToxCast was created as a priority setting and screening program. Assays were designed to maximize throughput, data processing was designed to minimize false negatives, and the data was intended to be used altogether for computational exercises and modeling to identify bioactivity patterns or profiles. However, the question is now often being asked "What does ToxCast say about chemical X?." The availability of data permits users to provide insight into this question, but the appropriate utilization of ToxCast data in this context requires deeper analysis and consideration of some key factors.

Curve/Data Quality

Due to the high-throughput nature of the ToxCast screening program and the large amounts of data collected in the program, automated data processing is absolutely necessary. Also, for a basic user of the data, it is often accessed only as a hit call (active or inactive) or as a hit call with the AC50. Simplifying the data to this level may be necessary for some computational exercises, but when considering the data for a single chemical, a more in-depth analysis of the data used to derive this hit call and AC50 is appropriate. Users should consider both the validity of the hit call and the plausibility of the AC50. The following paragraphs will present suggestions for a qualitative analysis of these parameters. More sophisticated, quantitative methods for improving or refining the hit call and AC50 determination processes are definitely possible but are not the focus of this document.

Not all responses that the data processing pipeline labels as "active" appear to be truly positive results. The ToxCast tcpl data processing package was created for screening and was therefore designed to minimize false negatives. False positives likely exist for most chemicals, and the end user is responsible for identifying these during their analysis. The first place to look to address the question of hit call accuracy is the dose-response curves provided in the Dashboard.

Exhibit 6: Example dose-response curves showing flag warnings.



A – Name of assay and chemical information can be found here

B – The flags and warnings noted for this data can be found here.

Source: https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID3023897#invitrodb-bioassays-toxcast-tox21

The ToxCast program has acknowledged that false positive and negative hit calls are possible using the automated methods and has thus added a processing step to assign "flags" or warnings to the data, as shown in Exhibit 6. While the flags are a helpful addition to the data analysis process, their assignment is also completely automated, and thus

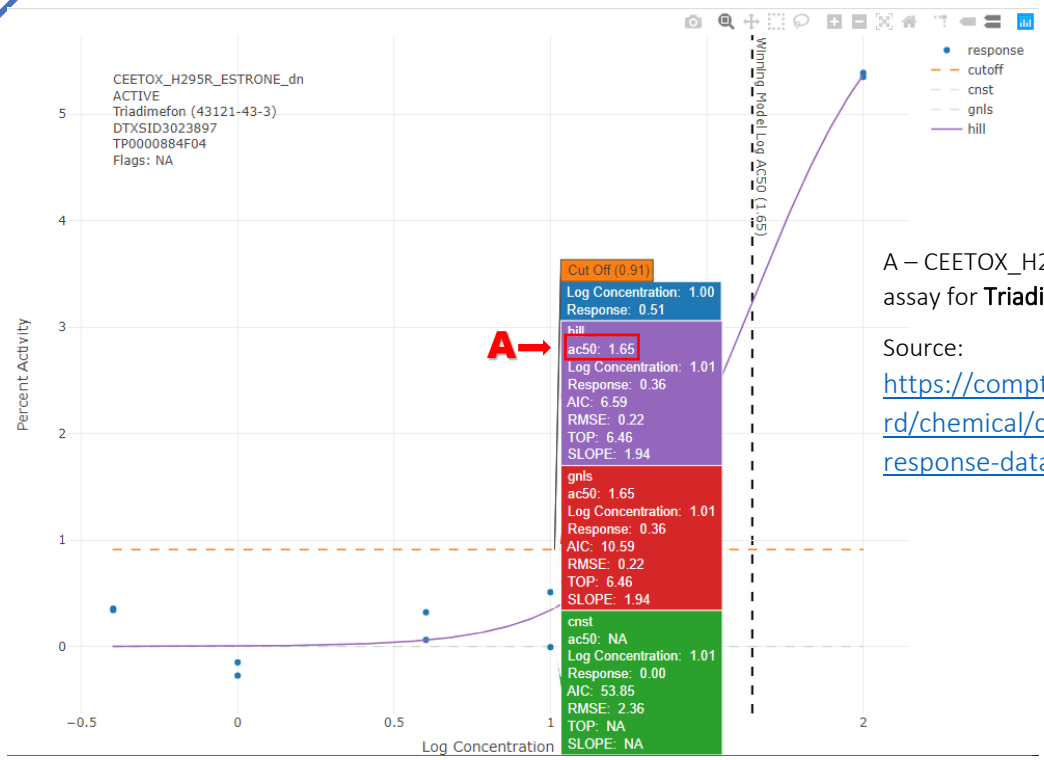
prone to some error. Careful examination of dose-response curves shows that within a group of curves receiving a certain flag, there is still a wide range of responses. Therefore, it may not be the best practice to set hard filters based on flags, but flags are definitely an important consideration when analyzing a list of positive results. After determining the accuracy of the hit call, it is also important to consider the AC50 that has been calculated for the response. Errors of this nature seem to be less frequent than false positive hit calls, but they still exist and should not be overlooked.

Comparison to Reference Chemicals

Comparison to reference chemicals is a common way to assess the validity of *in vitro* assays. Reference chemicals are identified for many ToxCast assays, but the high number of unique biological targets in ToxCast means that not every assay has a defined reference. Users may choose to identify their own reference chemicals relevant to their intended application of the data. Both the potency (AC50) and efficacy (top or max response) of the response of a chemical of interest can be compared to a reference chemical. Users can identify reference chemicals used for each assay within the Assay List (See Exhibit 1B). After selecting the desired assay, selecting the document icon  next to the assay name at the top of the screen will open that assay's details. Under the "Key Positive Controls" line you can find the reference chemicals used. These chemicals could then be accessed through the search bar and the assay results would be available (See Exhibit 2).

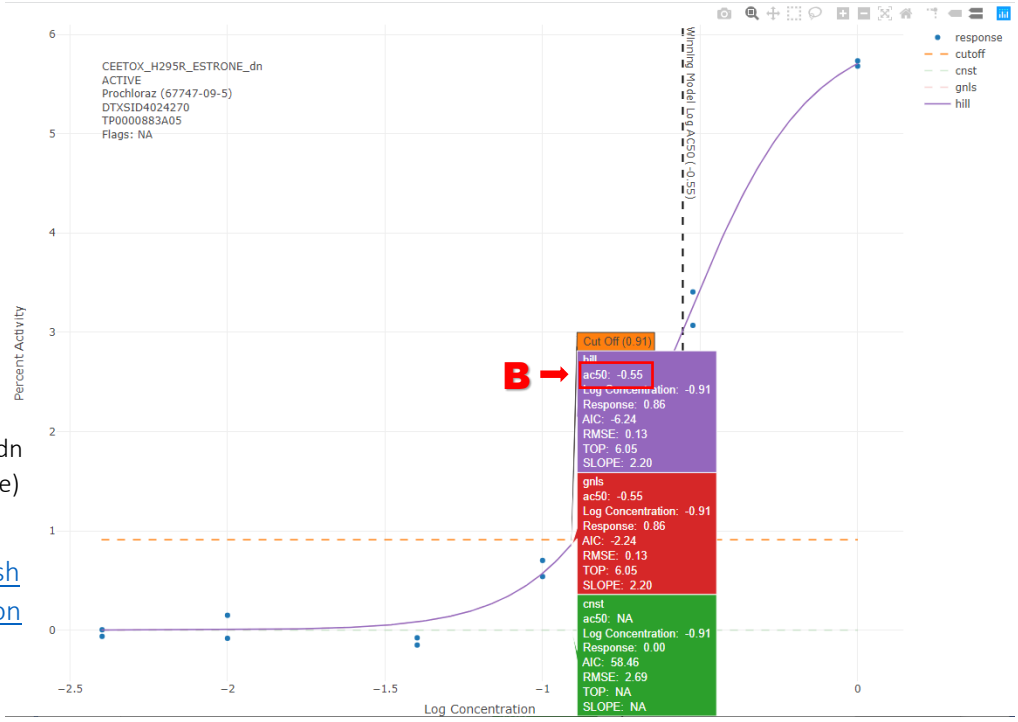
For example, as seen in Exhibit 7, the Ceetox assays for a decrease in estrone levels (CEETOX_H295R_ESTRONE_dn) use prochloraz as a control chemical. The curve for triadimefon looks similar to that for Prochloraz. The efficiency is also similar, both reaching approximately 60-fold induction (notice the log2 scale of the y-axis). However, comparing the AC50s, prochloraz is nearly 100-fold more potent than triadimefon (0.28 vs. 44.19 µM, respectively). This does not necessarily mean that the data for triadimefon is not relevant, as a range of potencies for different chemicals is expected, but this is one factor that should be considered, and final interpretation depends on the intended use of the data.

Exhibit 7: Comparison of the CEETOX_H295R_ESTRONE_dn assay for Prochloraz and Triadimefon Consistency Among Orthogonal Assays and/or Related Targets



A – CEETOX_H295R_ESTRONE_dn assay for **Triadimefon** (AC50 Value)

Source: https://comptox.epa.gov/dashboard/chemical/concentration-response-data/DTXSID3023897



B – CEETOX_H295R_ESTRONE_dn assay for **Prochloraz** (AC50 Value)

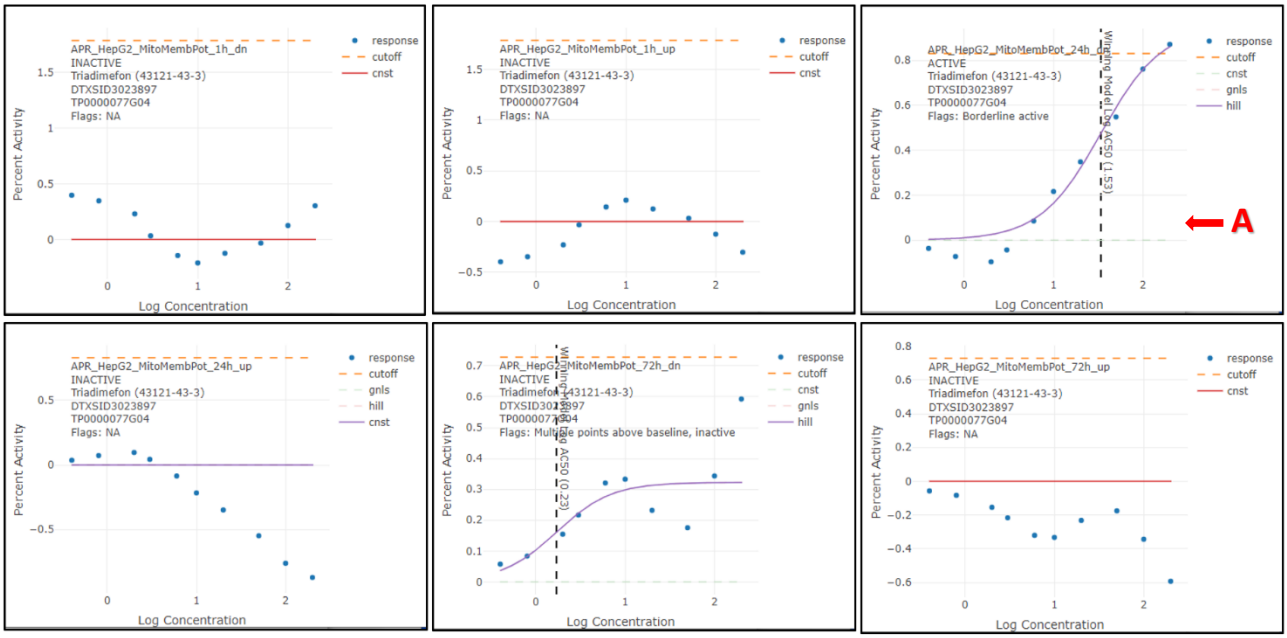
Source: https://comptox.epa.gov/dashboard/chemical/concentration-response-data/DTXSID3023897

As defined by Thorne et al, 2011, an orthogonal assay is an assay performed following the primary assay to differentiate between compounds that generate false positives from those compounds that are genuinely active against the target. Conducted on compounds found active in the primary assay, this assay uses a different reporter or assay format to confirm that activity of the compound is directed toward the biological target of interest. Compounds inactive in an orthogonal assay are removed from further consideration, as a negative result indicates that the original compound activity was most likely assay format-dependent and not specific to the biology of interest.

Several biological endpoints are covered by multiple ToxCast assays. Perfect consistency cannot be expected due to the high-throughput nature of the assays but comparing multiple assays can improve the interpretation and confidence in the data.

To illustrate a specific example using triadimefon, Exhibit 8 shows that there are six assays designed to measure mitochondrial membrane potential (MMP). In the Apredica (APR) technology HepG2 cells are used, and three time points are measured. The Tox21 assay measures the MMP in HepG2 cells at 24 hours. Triadimefon is only positive for one of these six assays, APR_HepG2_MitoMembPot_24h_dn. Even without an analysis of the dose-response curve for this assay, the lack of response in the other assays suggests that this may be a non-specific response. Overall, there is weak support for a specific disruption of mitochondrial membrane potential by triadimefon.

Exhibit 8: MMP Assays for Triadimefon



A – Active assay for Triadimefon: APR_HepG2_MitoMembPot_24h_dn

Source: https://comptox.epa.gov/dashboard/chemical/concentration-response-data/DTXSID3023897

Specificity

Specificity may be considered in the context of the promiscuity of an individual assay, or in the context of how many or which assays a single chemical hit. Information regarding the number of chemicals active for a given assay in relation to the total number of chemicals tested can be found in the Assay_Quality_Summary_Stats_190708.csv file accessed from this EPA file share as a zipped folder. The column test provides the number of samples tested, acnt provides the number of active samples and apct gives the percent active samples.

If one were to simply sort the assays based on the percent active samples (apct), the Novascreen (NVS) assays would clearly have the highest percent actives, up to 95%. However, this analysis is quite misleading as the NVS assays were first tested at a single concentration, and only actives from the initial screen were tested in multiple concentrations. The apct value is calculated on those chemicals selected for multiple concentration screening. Therefore, the high percent actives reflect the ability of the single concentration screen to identify positive chemicals rather than the true ratio of active chemicals out of all chemicals tested.

Another group of assays with a high number of actives is those related to PXR (ATG_PXRE_CIS_up and CLD gene expression assays for PXR target genes). Many of these assays are active for > 40% of chemicals tested. As PXR is a xenobiotic-sensing nuclear receptor with a wide variety of targets, the high percent actives may truly reflect the promiscuity of the PXR receptor rather than the ToxCast assays designed to measure its activity. Overall, there are no recommended guidelines for what percent active samples represent a “promiscuous” assay, and several factors must be considered, but this information may be helpful on an individual basis.

The other context for specificity is illustrated well by the Attagene (ATG) assays activated by triadimefon. Triadimefon activates 21 different ATG assays, which measure transcription factor activation (Figure 11 A). A closer look at the target of each assay shows that the ATG_CMV_CIS_up assay is a background control reporter used to indicate non-

specific interference. This indicates that it may not be possible to distinguish specific activity from this non-specific background activity for curves that resemble the activity in the CMV assay, such as that of the metal response element (MRE). The other positive ATG assays for triadimefon should be analyzed in a similar manner. Activation of the CMV background reporter assay does not necessarily mean that all results from ATG are unreliable, however the specificity of each response should be considered with greater scrutiny.

Q7. How can ToxCast data be applied inform decision making?

Replacing Animal Testing in a Tiered Testing Approach

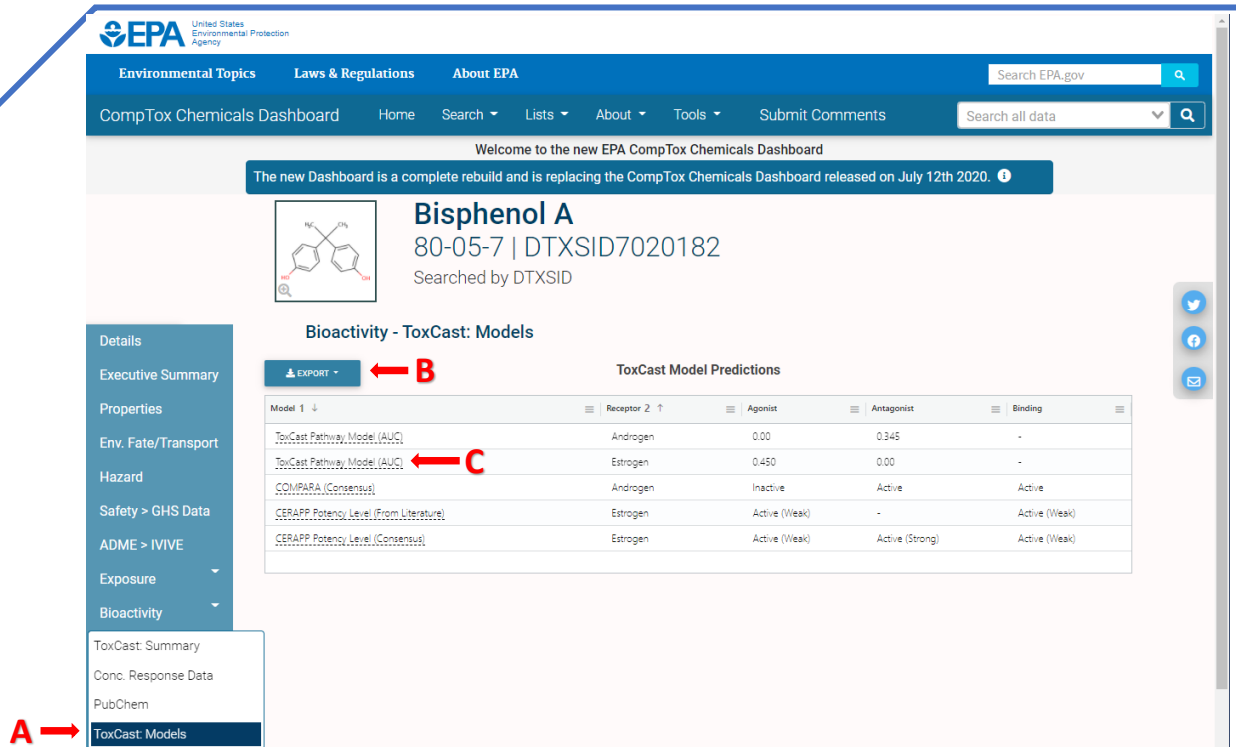
There are many ways in which ToxCast data can and is being used to inform decision making. While data for decision making has long revolved around *in vivo* experiments in various animal models, ToxCast data can be utilized in an early phase of a tiered approach in place of animal testing. In ToxCast's early days, EPA attempted to use ToxCast models to predict rodent endpoints, as exemplified in Thomas et al. 2016, A comprehensive statistical analysis of predicting in vivo hazard using high-throughput in vitro screening. These attempts resulted in predictions that were no better than chance at determining endpoints. Some examples of using ToxCast data in tiered testing approaches are listed below.

- In *Developing context appropriate toxicity testing approaches using new alternative methods (NAMs)*, Andersen et al discuss how a tiered approach to toxicity testing could yield more specific decision-making capabilities. The second tier of the system proposed involves high-throughput *in vitro* screening, such as the data that ToxCast can provide. It is, however, important to note that **while ToxCast data can be used in a tiered testing approach, it is not enough to make a prediction of adverse effects on its own.**
- As discussed previously, in Becker et al. 2015, *An exposure:activity profiling method for interpreting high-throughput screening data for estrogenic activity – Proof of concept*. the authors indicate how ER bioactivity can be evaluated using IVIVE and developing margins of exposure. Information such as this can then be used in a tiered testing approach to inform decisions.
- In *Hexabromocyclododecane (HBCD): A case study applying tiered testing for human health risk assessment*, a collaboration between the US EPA and Health Canada, bioactivity data was used for tiered testing. Using the bioactivity-exposure ratio (BER) the researchers looked to determine risk for HBCD. The tiers included IVIVE extrapolation of ToxCast data, BERs, and more.
- Another example is the US EPA's white paper *A Working Approach for Identifying Potential Candidate Chemicals for Prioritization* discusses various ways that ToxCast data can replace *in vivo* animal data.

The ToxCast Models subtab

Prediction models associated with bioactivity are all assembled under the ToxCast: Models subtab, as shown in Exhibit 9 below. For model details and links to papers, you can hover over the information icon.

Exhibit 9: ToxCast Models Subtab for BPA



A – ToxCast Models tab, located within the Bioactivity tab.

B – Export option to download the model data presented.

C – Hovering over the name provides more detail about the specific model.

Source: https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID7020182#bioactivity-toxcast-models

Q8. What are the limitations of ToxCast data?

Although ToxCast is an important screening tool along with other predictive tools and information (including exposure) to set priorities for chemicals with limited toxicity information, it is critical to be aware of ToxCast's limitations when interpreting results. One limitation of this data is that most of the chemicals tested are single substances not UVCBs (unknown or variable composition, complex reaction products or of biological materials) or mixtures, and wax-like substances and metals are not tested.

Interference of Cytotoxicity or Non-Specific Cell Stress

The ToxCast program recognizes that there is a need to distinguish between toxicity due to the disruption of a specific biomolecular function and toxicity due to general cell stress. This concept has been well described in a recent paper by Judson et al. *Analysis of the Effects of Cell Stress and Cytotoxicity on In Vitro Assay Activity Across a Diverse Chemical and Assay Space*. ToxCast has built in 86 assays to measure cytotoxicity in various cell types. The full list of these "burst" assays can be obtained by filtering the Assay_Summary_190708.csv file for 1's in the burst_assay column. This file can be downloaded from this EPA file share as a zipped folder. In the graphics from Judson et al., the median AC50 of the active burst assays (cytotox median) is shown in red, and confidence intervals around the median are calculated to give the lower limit (cytotox min), shown in the gray box. Any assays with positive results in the cytotoxicity region (gray box) should be carefully analyzed to determine if non-specific cytotoxicity may confound the activity. In *How well can carcinogenicity be predicted by high throughput "characteristics of carcinogens" mechanistic data?* by Becker et al., the burst phenomena is further discussed.

For many reasons, it may not be the best practice to use the cytotox limit as a hard filter. The median, and therefore lower limit are calculated based on active burst assays. As discussed in an earlier section, hit calls may not always be accurate, and the results for burst assays are determined in the same way as all other assays. Therefore, the median could be a misrepresentation of the true chemical response if one or more of the burst assays are potentially mislabeled as active (false positive). The cytotox min is based on a statistical calculation with uncertainties. The 86 burst assays do not represent all cell types used in the ToxCast assays, so it cannot be automatically assumed that all cell types would show cytotoxicity in the same concentration range. Although the method for identifying potential interference from cytotoxicity is not perfect, it can be quite effective when considered as a flag for greater scrutiny rather than a filter for relevant assay responses.

Limitations in the Interpretation of Bioactivity

The ToxCast assays are limited in terms of the biology they can recapitulate in an *in vitro* system. For example, most of the assays lack metabolic competency which means that the bioactivity of a chemical could be different when metabolized in a whole organism.

There are also scenarios in which the desired chemical concentration applied to an assay is not what reaches the target cells. Recognizing the potential difference between the nominal and realized concentrations is important when running assays and has caused some chemicals to not be included in the ToxCast data set due to their testing complexity. This reduction in the chemical diversity within ToxCast is a limitation of the entire data set. Some reasons why the nominal and realized concentrations may be different are listed below.

- The chemical may partition into the matrix of the assay.
- Volatile chemicals are difficult to stabilize properly to ensure that the desired concentration reaches the target cells.
- Chemicals that bind to plastics may influence results. This can be overcome using other materials, such as glass, in the testing, but must be taken into consideration.

There are other issues that may prevent a chemical from being run in the assay at all.

- Water insoluble chemicals are difficult to test in the current assays as their protocols require solubility.
- Surface active chemicals are difficult to test due to their preferential adsorption.

EPA has acknowledged that there are limitations in this area and have worked to improve the data. An example is that in the newest version of the dashboard, additional; assays conducted in HegRG cells are included and these have some metabolic activity. This area may see further improvement.

Q9. Where can I find additional resources about ToxCast?

There are many additional resources that may be helpful while using ToxCast data. Some good resources are listed below.

- Publications about the CompTox Dashboard and ToxCast
    - The CompTox Chemistry Dashboard: a community data resource for environmental chemistry
    - ToxCast Chemical Landscape: Paving the Road to 21st Century Toxicology

  - National Research Council's 2007 document: Toxicity Testing in the 21st Century: A Vision and a Strategy.

- EPA Webinars and Videos
  - EPA CompTox Dashboard Primer Videos
  - EPA CompTox Dashboard (ToxCast portion begins at 8:45 min)
  - ToxCast Owner's Manual Webinar
- EPA's ToxCast Owner's Manual
- NICEATM Resources
  - Link to NICEATM ICE User Guide
  - Link to NICEATM ICE FAQ
- Additional EPA Resources
  - Distributed Structure-Searchable Toxicity (DSSTox) Database