# Supplemental Methods

# Single-concentration (sc) screening analysis

## HCI assays (proliferation, apoptosis, NOG)

The goal of the single-concentration screen was to provide a health-protective estimate of which compounds would be a hit in the multiple-concentration screening, so in a first approach the hit-call determination was based on the same normalization methods used in the multiple-concentration screening in the ToxCast Pipeline (tcpl). For every HCI endpoint currently registered in tcpl (Table 1), the percent-of-control response values are calculated from the raw value (*rval*) and median DMSO *rval* on each plate (*bval)* with this formula:

Thus, the *resp.pc* represents a zero-centered, normalized response value. For all assay components except CCTE\_Mundy\_HCI\_Cortical\_NOG\_Casp3\_7, the *resp* value is multiplied by -1 so that a decrease in *resp* will lead to an increasing dose-response curve. The *bmad* is calculated as the median absolute deviation of the *resp.pc* in all control wells in the dataset. The current cutoff for HCI assay endpoints in multi-concentration screening in tcpl is defined as 3x*bmad* above the median control value or 30%, whichever is larger. (For 12 of the 17 assay endpoints currently in tcpl, 3x*bmad* is less than 30%).

**Table 1**: The 17 HCI endpoints in tcpl and the associated coff and 3xBMAD values for the multiple-concentration screening.

aeid aenm bmad3 coff

1: 2777 CCTE\_Mundy\_HCI\_Cortical\_NOG\_BPCount\_loss 23.578699 30.00000

2: 2778 CCTE\_Mundy\_HCI\_Cortical\_NOG\_NeuriteCount\_loss 4.601172 30.00000

3: 2779 CCTE\_Mundy\_HCI\_Cortical\_NOG\_NeuriteLength\_loss 23.136021 30.00000

4: 2780 CCTE\_Mundy\_HCI\_Cortical\_NOG\_NeuronCount\_loss 22.087479 30.00000

5: 2781 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_BPCount\_loss 17.737887 30.00000

6: 2782 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_CellBodySpotCount\_loss 44.678480 44.67848

7: 2783 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuriteCount\_loss 7.990187 30.00000

8: 2784 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuriteLength\_loss 22.795974 30.00000

9: 2785 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuriteSpotCountPerNeuriteLength\_loss 24.710000 30.00000

10: 2786 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuriteSpotCountPerNeuron\_loss 36.575546 36.57555

11: 2787 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuronCount\_loss 16.039668 30.00000

12: 2788 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_SynapseCount\_loss 31.165243 31.16524

13: 2793 CCTE\_Mundy\_HCI\_hNP1\_Casp3\_7\_gain 9.269828 30.00000

14: 2794 CCTE\_Mundy\_HCI\_hNP1\_CellTiter\_loss 5.874010 30.00000

15: 2795 CCTE\_Mundy\_HCI\_hNP1\_Pro\_MeanAvgInten\_loss 31.030868 31.03087

16: 2796 CCTE\_Mundy\_HCI\_hNP1\_Pro\_ObjectCount\_loss 23.183304 30.00000

17: 2797 CCTE\_Mundy\_HCI\_hNP1\_Pro\_ResponderAvgInten\_loss 34.292636 34.29264

**Verifying the Proposed hit call method**

The available multiple-concentration HCI data in tcpl were analyzed to verify if the median *resp.pc* at the highest concentration tested could provide a health-protective metric for the hit-call in a single-concentration screen against 2 or 3 x *bmad*. Particularly, the hits where the median *resp.pc* is less than the *max\_med* were examined(e.g. if the dose-response curve is “biphasic” (aka ‘gain-loss’) or if the dose response curve is slightly noisy). The available HCI data in tcpl includes ~80+ compounds.

The 2x*bmad* threshold was found to detect all but 4 compounds that were labelled as “hits” in the multiple-concentration screening. The 4 compounds that were missed all fit a “gain-loss” dose-response curve.

hitc aeid aenm spid bmad med\_resp\_at\_max\_conc

1: 1 2780 CCTE\_Mundy\_HCI\_Cortical\_NOG\_NeuronCount\_loss EX000389 7.362493 -3.0303030

2: 1 2785 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuriteSpotCountPerNeuriteLength\_loss TT0000177C03 8.236667 -11.1111111

3: 1 2787 CCTE\_Mundy\_HCI\_Cortical\_Synap&Neur\_Matur\_NeuronCount\_loss EX000402 5.346556 0.6858123

4: 1 2793 CCTE\_Mundy\_HCI\_hNP1\_Casp3\_7\_gain TT0000177E03 3.089943 -1.6081565

Conclusion: 2x*bmad* will provide a good estimate for the single-concentration screen hit calls.

**Applying Single-Concentration Hit Calls to PFAS data**

* Pre-processing
  + Removed any data points where the Well quality is 0
  + Renamed any wells labelled as “Blank”, “NA”, or “Media” from each assay as “Media”
  + Defined the “wllt” (well type) as:
    - If Compund.Name is Media or DMSO, wllt = “n” to signify neutral controls. (In contrast, the tcpl single-concentration methods only utilize DMSO wells).
    - Otherwise, wllt = “t” (to signify test compounds)
* Found the median raw value of control wells on each plate (where wllt == “n”) (*bval*)
* Calculated the normalized, zero-centered *resp.pc* from every raw value
* Calculated the rescaled Median Absolute Deviation of wells where wllt == “n” for each endpoint:
* Found the median *resp.pc* for each sample and endpoint
  + In order to be more health-protective, where there were only 2 replicates with Well Quality=1, set the *med\_resp\_up* as the maximum *resp.pc*, and the *med\_resp\_dn* as the minimum *resp.pc*
* Determined the hit calls for each sample for each endpoint:
  + If the median *resp.pc\_dn* is less than or equal to -2x*bmad*, then the sample is a “down” hit
  + If the median *resp.pc\_up* is greater than or equal to 2x*bmad*, then the sample is an “up” hit
  + Otherwise, the sample is a “no hit”

**Final results**

There are 53 samples with a hit in at least 1 assay (Supplemental Methods Table 1). See ‘hit\_call\_summary’ sheet for summary of hit-call results for any HCI assay (‘any\_hit’ column). For raw data values for this analysis, see sheets labeled ‘all\_values’ or ‘values\_for\_no\_hit\_compounds’ for data on compounds with activity or no activity, respectively.

## MEA NFA

This work aims to find an appropriate combination of endpoints and thresholds to determine which compounds tested in the Network Formation Assay (NFA) single-concentration screen are likely to be active in the NFA multi-concentration screen. The response values at the highest concentration tested and corresponding hit calls from the multi-concentration screen were used to select the most informative endpoints and to develop cutoffs. Then, the cutoffs for the selected set of endpoints were applied to the median percent-of-control response values of the PFAS NFA single-concentration screen to determine which compounds should be re-tested in the multi-concentration screen.

**Multi-concentration Analysis**

Preparation of multi-concentration data

The current multi-concentration NFA data set includes 422 samples, taken from several chemical sets, with the prefix “CCTE\_Shafer\_MEA\_dev\_.” Data were processed as follows:

* Removed any hits from the multi-conc data with 3 or more flags.
* Removed any hits where the AC50 is less than the lowest concentration tested and the top modl parameter is less than 1.2\*cutoff (fit category 36 or 45).

A sample was defined as “positive” in the multi-concentration screen if it has 3 or more unfiltered hits. This resulted in 236 “positives” and 186 “negatives.”

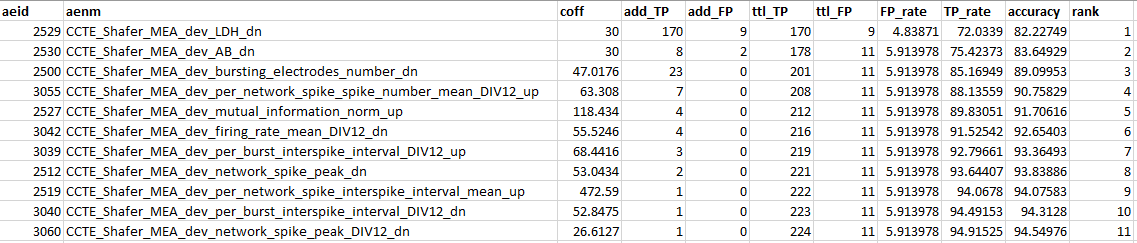
Selection of Endpoints and Cutoffs

The goal of this analysis is to identify a combination of endpoints and cutoffs that will detect as many positives as possible while minimizing the number of false positives. There are 66 total endpoints to choose from: 2 cytotoxicity endpoints (LDH and AB), 17 parameters analyzed in the up and down direction, and 15 “DIV12” endpoints analyzed in the up and down direction (2 + 17\*2 + 15\*2 = 2+34+30 = 66 total).

Algorithm to compile a set of endpoints and cutoffs (based on the “greedy algorithm”):

1. Start with the 2 cytotoxicity endpoints (LDH and Alamar Blue) as the initial endpoints with a cutoff of 30% for each (corresponding to a 30% decrease in viability relative to controls). Any sample with a median response at the highest concentration tested above 30 for either endpoint is labelled as “sc\_positive.” This resulted in the detection of 178 true positives and 11 false negatives.
2. For each remaining endpoint, set the cutoff just above the highest response for the remaining undetected negatives. Specifically,
   1. Set the lower bound of the cutoff as the highest med\_resp\_max\_conc of the undetected negatives for each endpoint
   2. Set the upper bound of the cutoff as the lowest med\_resp\_max\_conc of the undetected positives that is above the lower bound of the cutoff
   3. Set the cutoff as the mid-point between the upper and lower bounds. In this way, the cutoff is high enough that no additional negatives will be a “hit.”
3. Add the endpoint to the set that will add the most additional true positives. If there is a tie among endpoints, arbitrarily select one of them. Any sample with a median response at the highest concentration tested above the cutoff for the selected endpoint is labelled as “sc\_positive.”
4. Repeat from 2) until no additional true positives can be detected.

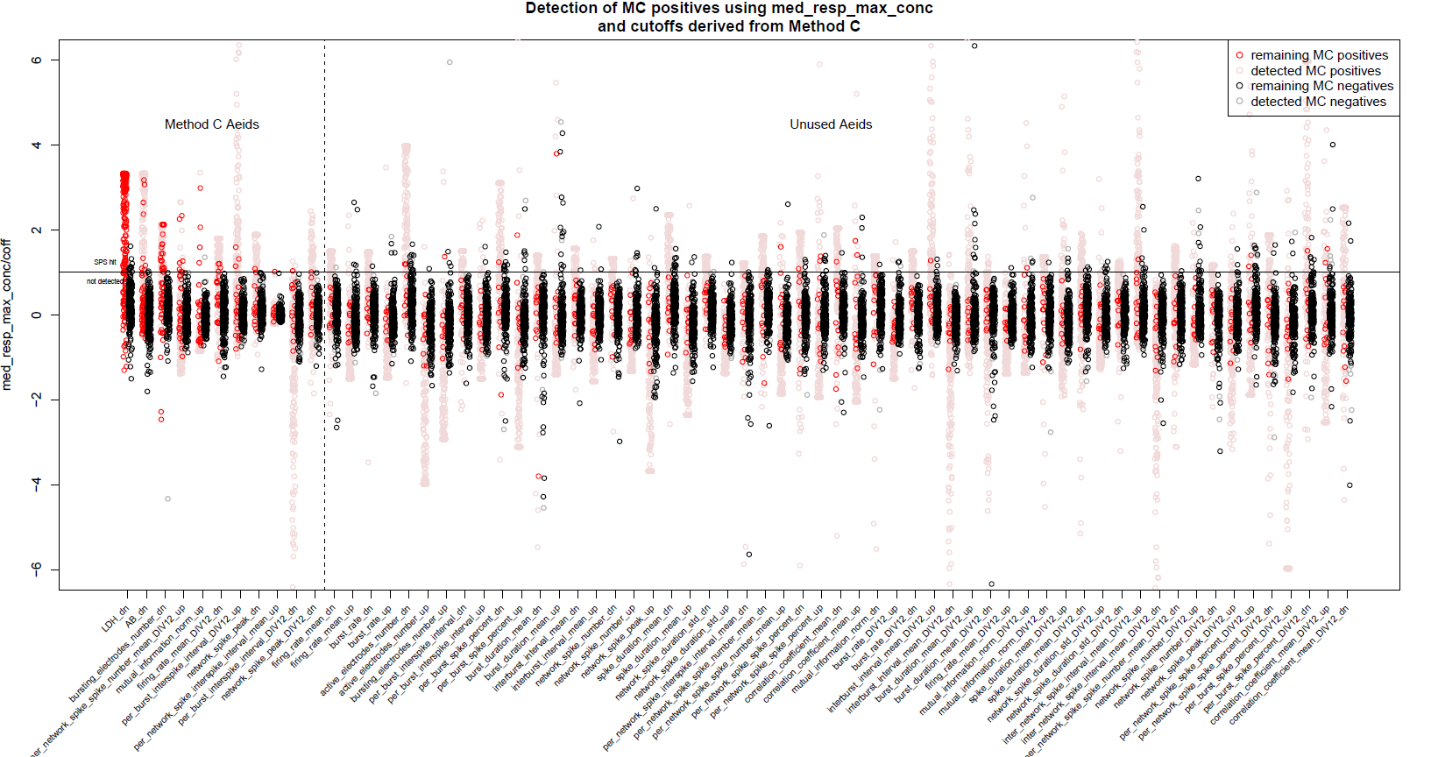
The algorithm was run at every juncture where there was a tie in the number of additional true positives at step 3. This resulted in 3 unique sets of endpoints and cutoffs (Supplemental Methods Table 2). One of the 3 sets of endpoints with the corresponding cutoffs is shown below. Each row shows the accuracy of the detection of the multi-concentration positives after the addition of each endpoint.



The other 2 sets of endpoints varied by only 1 endpoint. They contained either "CCTE\_Shafer\_MEA\_dev\_mutual\_information\_norm\_DIV12\_up" or "CCTE\_Shafer\_MEA\_dev\_per\_burst\_spike\_percent\_DIV12\_up" in place of “CCTE\_Shafer\_MEA\_dev\_mutual\_information\_norm\_up.”

Using any of the 3 sets of endpoints and the given cutoffs, 224 of the 236 positives could be detected with only 11 false positives out of the 186 negatives. (The true positives and false positives detected were identical with each of the 3 sets). This resulted in a final accuracy of accuracy of 94.55%. Interestingly, the majority of the true positives were detected by the LDH and Alamar Blue endpoints.

Below is a visual depiction of the detection of the multi-concentration positives and negatives (Figure 1). The endpoints are sorted by order of addition to the set.



*Figure1. The plot shows that the responses in the 12 remaining undetected positives are largely not separable from the negatives. See Supplemental Methods Table 2 for 3 unique sets of endpoints. The y-axis shows the median response at the highest conc tested for each sample, divided by the cutoff. Thus, any point above y=1 for the first 11 endpoints is a “hit” in the single-concentration screen. An arbitrary cutoff of 3\*BMAD was used for the remaining 55 endpoints.*

Consideration of 3\*BMAD for Cytotoxicity Cutoffs

The choice of 30% for the LDH and Alamar Blue cutoffs was somewhat arbitrary. In the multi-concentration screen in tcpl, the cutoff is set to 3 times the median absolute deviation of the percent-of-control values in DMSO wells (BMAD). This corresponds to a cutoff of 24% for LDH and 20% for Alamar Blue. The algorithm was run using 3\*BMAD as the cutoffs for LDH and Alamar Blue. This resulted in 21 false positives and 227 true positives. The addition of 3 true positives at the cost of 10 additional false positives did not seem beneficial. Therefore, 30% as the cutoff for the 2 cytotoxicity endpoints.

**Application of the Endpoints and Cutoffs to the PFAS data**

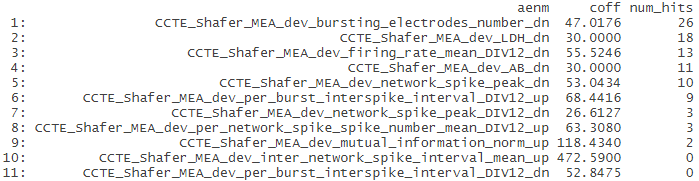
Normalization and Hit Call Determination

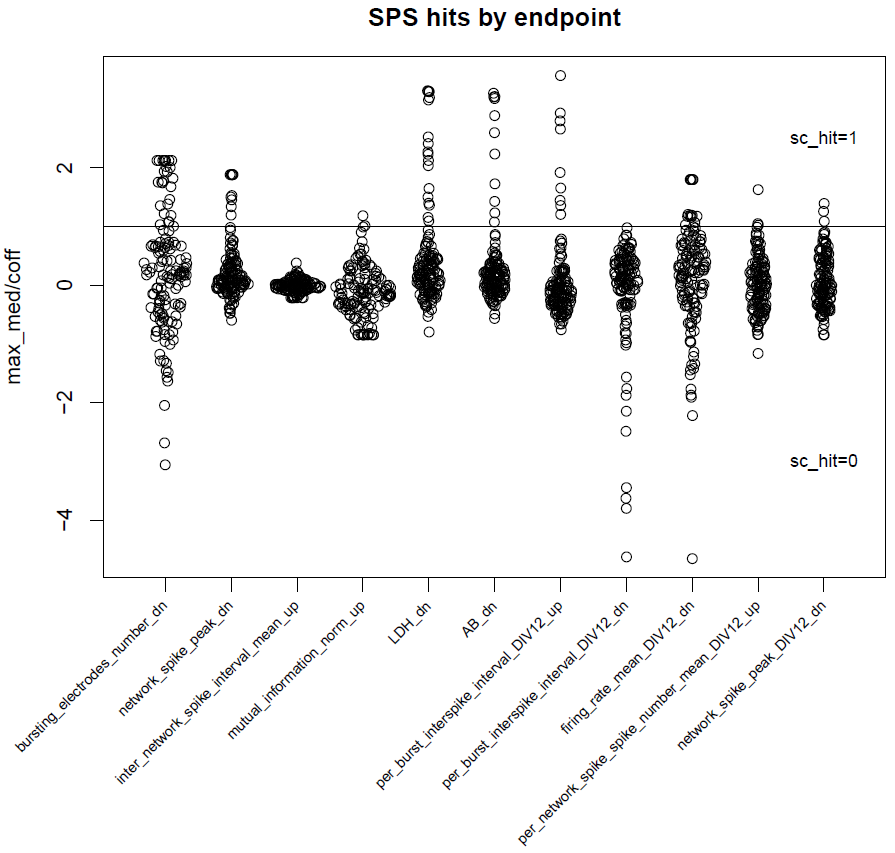
Single-concentration screen data processing:

* Calculated the “bval” as the median endpoint value of the DMSO control wells from each MEA plate
* Calculated the normalized “response” values:
  + To measure the “up” response:
  + To measure the “down” response:
* Calculated the median response value of each endpoint for every sample
* Determined the hit calls using the endpoints selected in the multi-concentration analysis
  + If the median response value for a given sample was greater than or equal to the cutoff for any of the 11 endpoints in the set, then that sample was labelled as a “positive”
  + Otherwise, it was labelled as a “negative”

The hit-call determination was repeated with each of the 3 unique sets of endpoints found in the previous section. The first set resulted in 36 positives (including the positive control Bisphenol). The second set resulted in the same 36 positives plus 1 additional sample. The third set resulted in the same 37 positives as the previous sets, plus 1 additional sample. Since the third set resulted in the most hits, the third set was used to assign the final hit calls (Figure 2).

The number of hits per endpoint are summarized in the table and graph below:





*Figure 2: Each point corresponds to the median response value for a sample divided by the cutoff. Any point above the horizontal line where max\_med/coff = 1 is a “positive.”*

**Final results**

Positive/negative determination for each sample and hit-calls for each sample for each of the 11 endpoints (“sc\_hit”) can be found in Supplemental Methods Table 3. Cutoffs were set to the cutoffs determined in the multi-concentration analysis.