**Network Formation Assay Single Point Screen Hit Call Criteria**

Feb 23, 2021

Amy Carpenter

**Summary**

This work aims to find an appropriate combination of endpoints and thresholds to determine which compounds tested in the Network Formation Assay (NFA) single point 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 point screen (SPS) to determine which compounds should be re-tested in the multi-concentration screen.

**Multi-concentration Analysis**

Script: *nfa\_sps\_endpoint\_cutoff\_selection\_2021-01-15.R, greedy\_algo\_methodC\_3.R*

Preparation of multi-concentration data

The current multi-concentration NFA data set includes 422 samples, taken from several chemical sets. These include DNT, ToxCast, NTP91, PFAS (76 compounds, tested in 2018), organophosphates, and 86 compounds tested in Frank et al., 2017.

I took the level 5 and 6 data from TCPL for the endpoints with the prefix “CCTE\_Shafer\_MEA\_dev\_.” Then, I processed the data as follows:

* I removed any hits from the multi-conc data with 3 or more flags.
* I 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).
* I included all AUC and DIV12 endpoints except for “active\_electrodes\_number\_DIV12\_up/dn” and “bursting\_electrodes\_number\_DIV12\_up/dn.” (I excluded these endpoints because the cutoffs are currently 0, making the hit calls extremely promiscuous).

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

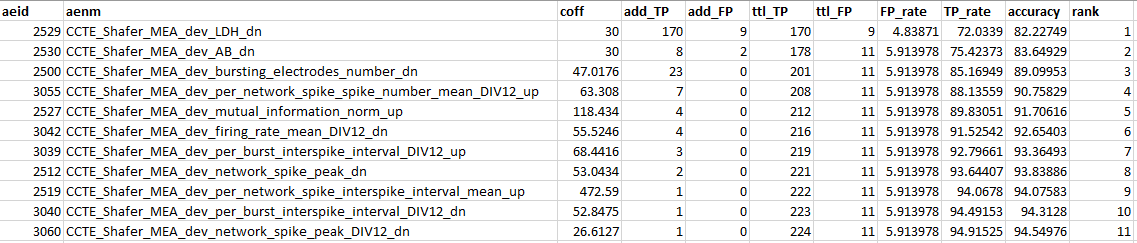
Selection of Endpoints and Cutoffs

I wanted to find 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 AUC components analyzed in the up and down direction, and 15 DIV 12 endpoints analyzed in the up and down direction (2 + 17\*2 + 15\*2 = 2+34+30 = 66 total). I hypothesized that some endpoints are much more useful than others for separating true positives from true negatives using only the highest concentration tested. I wanted to find those endpoints and appropriate cutoffs.

I used the following 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.

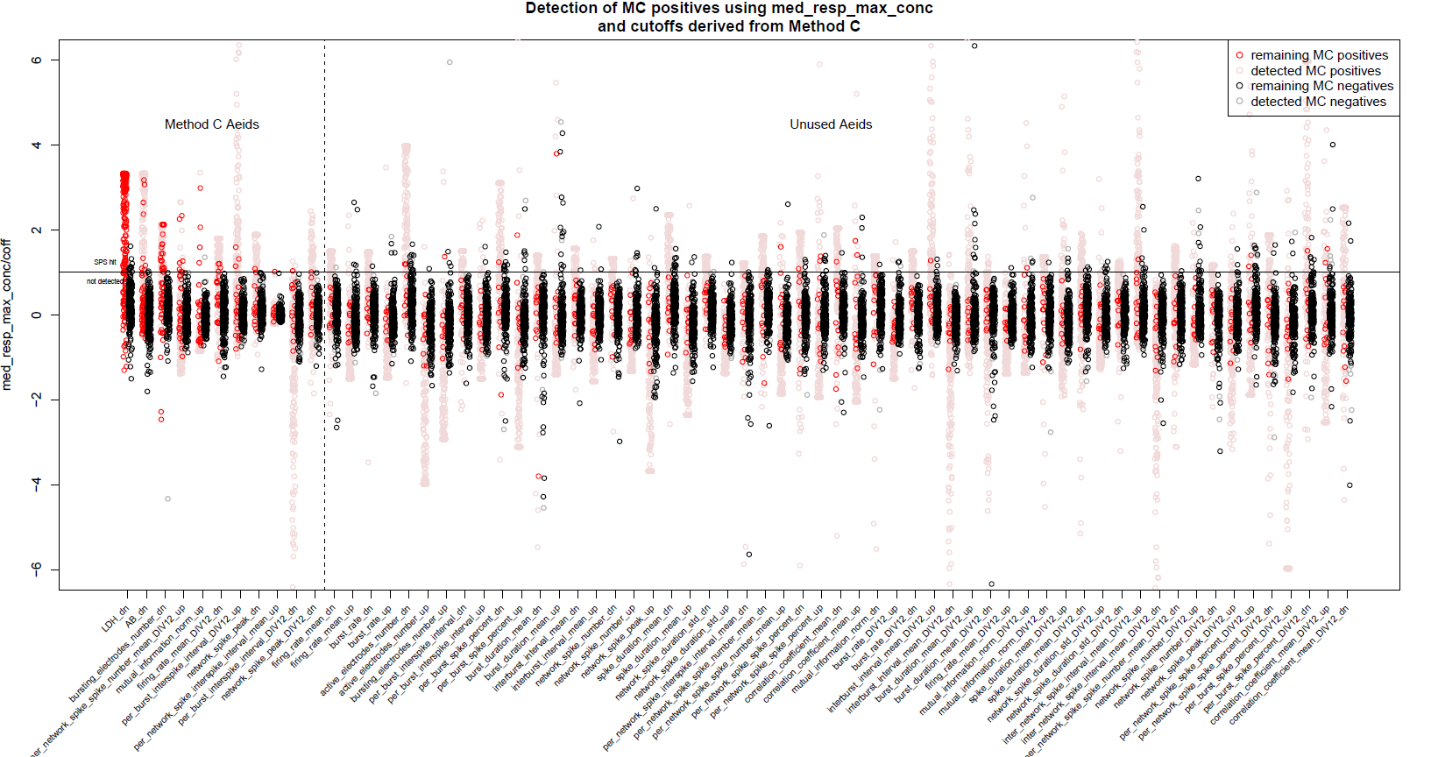
I re-ran the algorithm at every juncture where there was a tie in the number of additional true positives at step 3. This resulted in 3 unique sets of endpoints and cutoffs. One of the 3 sets of endpoints with the corresponding cutoffs is shown below. Each row shows the status after the addition of each 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. The endpoints are sorted by order of addition to the set.



*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 point screen. An arbitrary cutoff of 3\*BMAD was used for the remaining 55 endpoints.*

The figure shows that the responses in the 12 remaining undetected positives are largely not separable from the negatives. I believe that this model is fairly well optimized to determine which compounds are likely to be active in 3 or more multi-concentration endpoints using only the highest concentration tested.

The 3 unique sets of endpoints are saved under “endset\_tables\_2021-03-19.xlsx” and “endset\_tables\_2021-01-25.RData.”

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, 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. I ran the algorithm 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, I decided to use 30% as the cutoff for the 2 cytotoxicity endpoints.

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

PFAS NFA SPS MEA Raw Data Preparation

Scripts: *L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\nfa-spike-list-to-mc0-r-scripts\R\run\_me\_SPS\_PFAS2019.R*

Project folder: *L:\Lab\NHEERL\_MEA\Project PFAS 2019\MEA NFA*

I collected all of the spike list files, maestro experimental log files, and Calculations files from the following sub folders of the project folder:

* 20201118 Culture SPS G1
* 20201209 Culture SPS G2
* 20201209 Culture SPS G3
* 20201209 Culture SPS G4

I calculated the feature values (such as mean firing rate, burst rate, etc.) and the AUC values using the same scripts and procedures used to prepare the multi-concentration NFA data (see *L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\nfa-spike-list-to-mc0-r-scripts\R*). These scripts rely heavily upon the *meadq* and *sjemea* packages. Additional notes on the data processing can be found in the following notebook: *L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\Pre-processing\_MEA\_NFA\_for\_TCPL* (see tab “SPS\_PFAS2019”).

The output from the scripts can be found here: *L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\SPS\_PFAS2019.* The final output of the scripts is saved as *dat\_SPS\_PFAS2019\_2021-01-06.RData*.

Normalization and Hit Call Determination

Script: *nfa\_sps\_pfas2019\_hitcall\_assignment\_2021-01-25.R*

I processed the single point screen data as follows:

* Loaded all AUC and DIV12 endpoint values for the PFAS NFA SPS
* Removed all of the LDH points from plate MW72-8203, culture date 20201209. This plate was determined to be off, particularly in the control wells. See graph bellowing comparing the corrected optical density values from each plate in the data set.

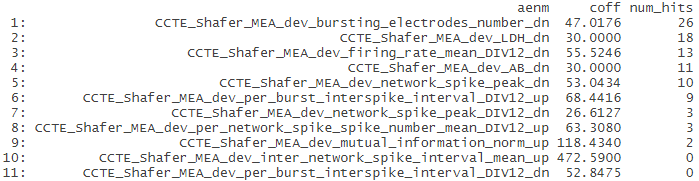
Chart

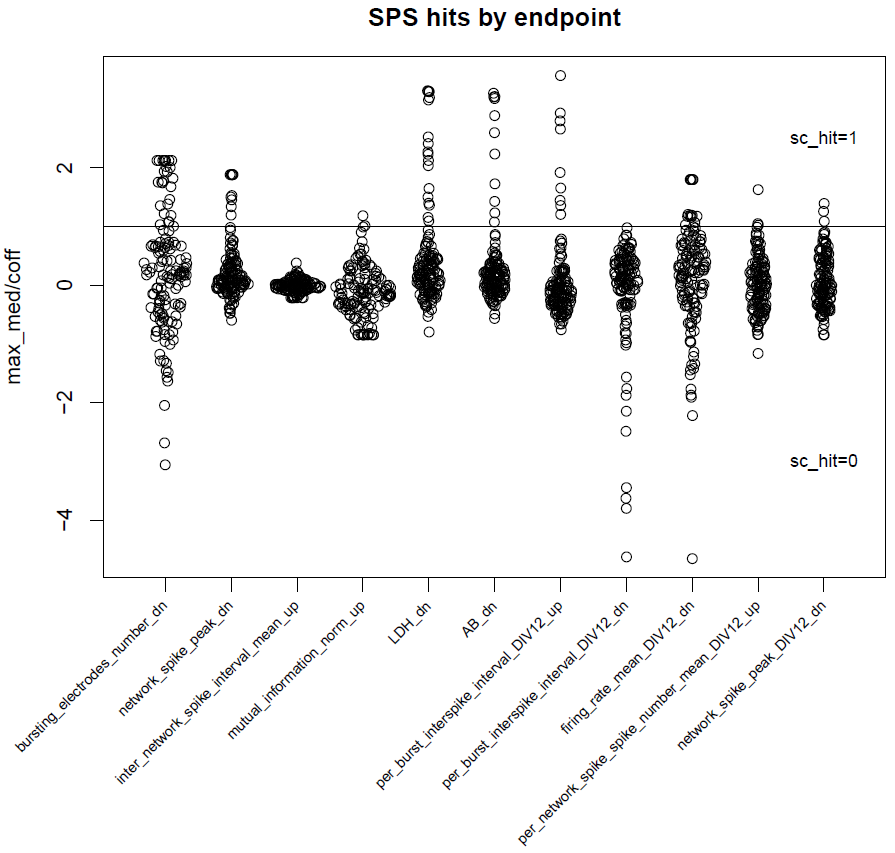
Description automatically generated

* 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”

I repeated the hit call determination 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, I used the third set to assign the final hit calls.

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





*Figure: 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.” Plot created with R/sinaplot*

Consideration of setting the cutoffs as a multiple of BMAD

I considered assigning the cutoffs for each endpoint as the median absolute deviation (MAD) of the response values of the DMSO control wells (BMAD) times the multipliers found in the multi-concentration analysis. However, I found that the BMAD’s for the PFAS SPS are almost exclusively larger than the BMAD’s of the multi-concentration data set. The difference in the BMAD’s is probably due to the difference in size/timespan of the data sets. (For the PFAS single point screen data set, there are 54-60 control points per endpoint, collected within 2 months in 2020. For the multi-concentration data set, there are ~1,500 control points per endpoint, collected from 2014-2019). Therefore, the SPS BMAD\*bmad\_multiplier would produce higher cutoffs than the multi-concentration BMAD\*bmad\_multiplier for almost all endpoints. So, in order to be more permissive in the hit calls, I used the multi-concentration BMAD\*bmad\_multiplier as the cutoffs for the single point screen hit calls.

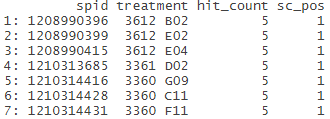
Chart, scatter chart

Description automatically generated

*Figure: BMAD’s from SPS data set are greater than BMAD’s from multi-concentration data set for most endpoints. The 10 endpoints used in all 3 sets of endpoints are shown in blue. The endpoints that vary among the 3 sets are in green. Unused endpoints are in gray. BMAD from the SPS was less than the BMAD from the MC for 2 endpoints: Alamar blue (6.75 in bmad\_mc, 6.21 in bmad\_sps) and Network spike peak (9.06 in bmad\_mc, 7.61 in bmad\_sps).*

Consideration of NA Raw Values

Seven samples had a value of “NA” for some endpoints. Thus, I could not compare the median response value to the cutoffs for these endpoints for the 7 samples. I considered labelling these samples as a “positive” to error on the side of inclusion. However, I found that all 7 of these samples were a hit under 5 other endpoints. Therefore, these samples were a labelled as “positive” regardless of the NA values.



Where to view results

Positive/negative determination for each sample (see “sc\_pos” column): *pfas\_sps\_hits\_2021-01-25.csv*

Hit calls for each sample for each of the 11 endpoints (“sc\_hit”): *pfas\_sps\_hits\_with\_aenm\_2021-01-25.csv*

Note that there were 5 PFAS samples that were not tested in the single point screen. These 5 will be added to the list of samples to re-test in the multi-concentration screen.

**Script summary**

*nfa\_sps\_endpoint\_cutoff\_selection\_2021-01-15.R, greedy\_algo\_methodC\_3\_list.R*

Description:

* Determines the desired combination of endpoints and cutoffs to use for the single point screen based on existing multi-conc NFA data

Output:

* endset\_tables\_2021-01-25.RData
* endset\_tables\_2021-03-19.xlsx
* endpoint\_set\_example\_with\_mc\_dat\_2021-03-19.xlsx
* nfa\_sps\_coffs\_visualization\_all\_aeids\_2021-01-25.pdf
* nfa\_sps\_coffs\_visualization\_selected\_aeids\_2021-01-25.pdf

*L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\SPS\_PFAS2019/run\_me\_SPS\_PFAS2019\_2021-01-06.R*

Description:

* Calculates the endpoint values from the raw MEA data

Output:

* dat\_SPS\_PFAS2019\_2021-01-06.RData
* (see *L:\Lab\NHEERL\_MEA\Carpenter\_Amy\pre-process\_mea\_nfa\_for\_tcpl\SPS\_PFAS2019/output\_2021-01-06* for additional output)

*nfa\_sps\_pfas2019\_hitcall\_assignmnet\_2021-01-25.R*

Description:

* Processes the PFAS SPS mc0 data, finds max\_med (median resp values) from each sample/endpoint pair
* Assigns the hit calls based on selected endpoints and cutoffs

Output:

* bmad\_comparison\_sps\_mc\_2021-01-30.png
* pfas\_sps\_hits\_2021-01-14.csv
* pfas\_sps\_hits\_with\_aenm\_2021-01-14.csv
* SPS\_PFAS2019\_normalized\_dat\_2021-01-25.RData
* SPS\_PFAS2019\_hit\_calls\_2021-01-25.RData
* sps\_hits\_by\_endpoints2021-01-25.pdf

**Future directions**

There is a lot that could be done to improve the selection of endpoints and cutoffs. But, a 95% sensitivity seems acceptable to me (though I do want to verify the sensitivity with an overfit analysis). Here are a few specific ideas of ways to improve and/or verify the methods:

* Do an overfit analysis to get a more robust assessment of the sensitivity/trustworthiness of the multi-conc analysis
* Consider lowering the cutoffs and allowing a few additional false negatives (in order to detect the 12 false negatives from the multi-conc analysis)
* Look at the distribution of values at the highest concentration tested to get confidence intervals, rather than just the median value. Then use that information to determine which endpoints have the best separation between positives and negatives at the highest concentration tested.
* Use Recursive Feature Elimination to find the most informative endpoints (rather than just selecting the endpoints that can contribute additional true positives without additional false positives)