OECD Meeting of the expert group on DNT IVB

Organisation for Economic Co-operation and Development

Item 5.c. Chemicals Tested in the DNT IVB Assays, analysis of data, and mathematical models

Kelly Carstens, PhD

U.S. Environmental Protection Agency
Office of Research and Devleopment
Center for Computational Toxicoogy and Exposure
Research Triangle Park, NC

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.

April 11, 2024

Organisation for Economic Co-operation and Development

ENV/CBC/TG/A(2024)3

For Official Use English - Or. English

27 March 2024

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

Cancels & replaces the same document of 27 March 2024

Working Party of National Coordinators of the Test Guidelines Programme

Draft Agenda of the Meeting of the Expert Group on Developmental Neurotoxicity (DNT)-In Vitro Battery (IVB)

11 April 2024
University of Konstanz, Konstanz, Germany

Item 5.c. Chemicals Tested in the DNT IVB Assays, analysis of data, and mathematical models

Updates on Chemicals Tested in the DNT IVB Assays, data availability and update of Appendix E.

Sharing methodologies to analyse data and ways to come up with one interpretation.

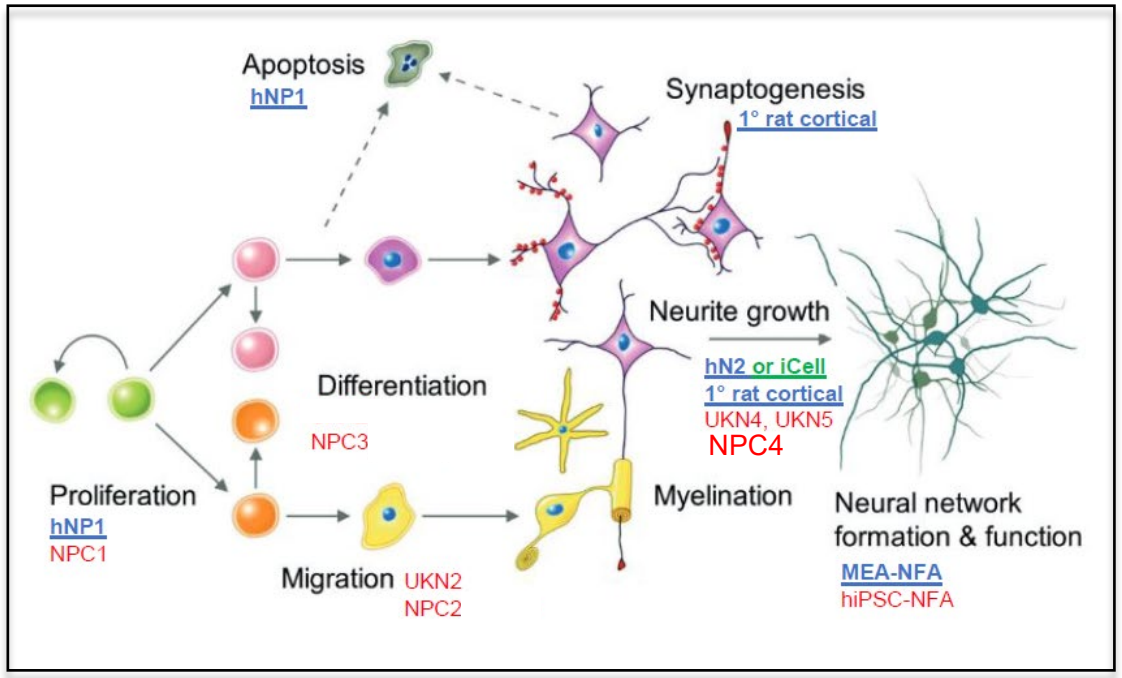
Advances analysis and integration of the existing DNT-IVB data and opportunities to

- Understand and explain differences in data from assays that measure the same developmental process.
- Understand the sensitivity of some methods in the battery and advise for which purpose they should be used or not.
- Use machine learning to solve the problem of integrating data from the battery.

Opportunity to share advancements in mathematical models and considerations about developing a Defined Approach.

Updates on Chemicals Tested in the DNT IVB Assays, data availability and update of Appendix E.

Data Landscape: DNT IVB assays currently in the ToxCast database*

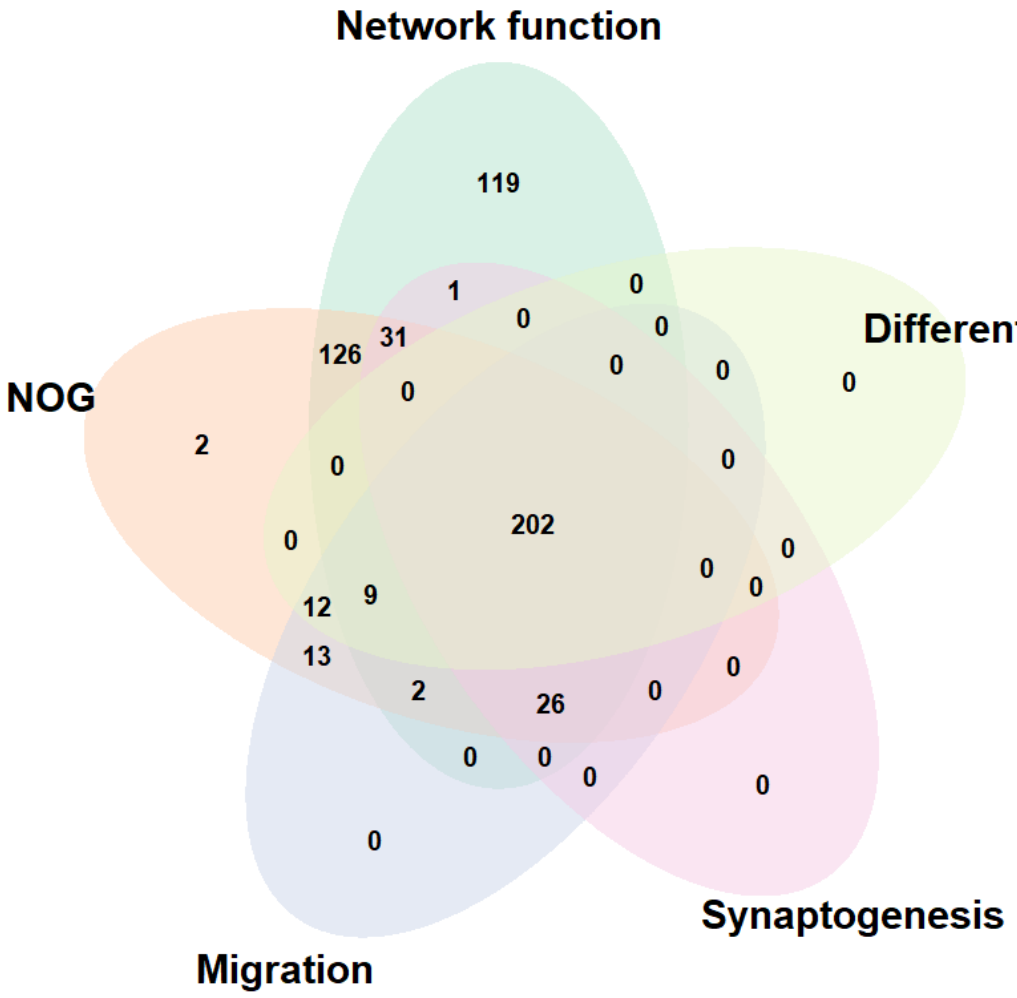


https://www.regulations.gov/document/EPA-HQ-OPP-2020-0263-0054

	Neurodevelopmental Process	Species	Assay	N chem tested
1	Apoptosis	Human	Apoptosis/viability, hNP1	385
2	Differentiation	Human	NPC3, Differentiation, neuron	224
3	Differentiation	Human	NPC5, Differentiation, oligo	224
4	Migration	Human	UKN2	237
5	Migration	Human	NPC2, Migration, radial glial	224
6	Migration	Human	NPC2, Migration, neuron	224
7	Migration	Human	NPC2, Migration, oligo	224
8	Network function	Rat	MEA NFA	517
9	NOG	Human	UKN5	146
10	NOG	Human	UKN4	144
11	NOG	Rat	NOG, rat	263
12	NOG	Human	NOG, human hN2	85
13	NOG	Human	NPC4, NOG	224
14	NOG	Human	NOG, CDI iGluta	309
15	NOG	Human	NOG, CDI iGABA	28
16	Proliferation	Human	Proliferation, NPC1	224
17	Proliferation	Human	Proliferation, hNP1	386
18	Synaptogenesis	Rat	Synaptogenesis/maturation, rat	261

*ToxCast invitrodb (release summer 2024)

Data Landscape: Overlap of chemicals tested across assays

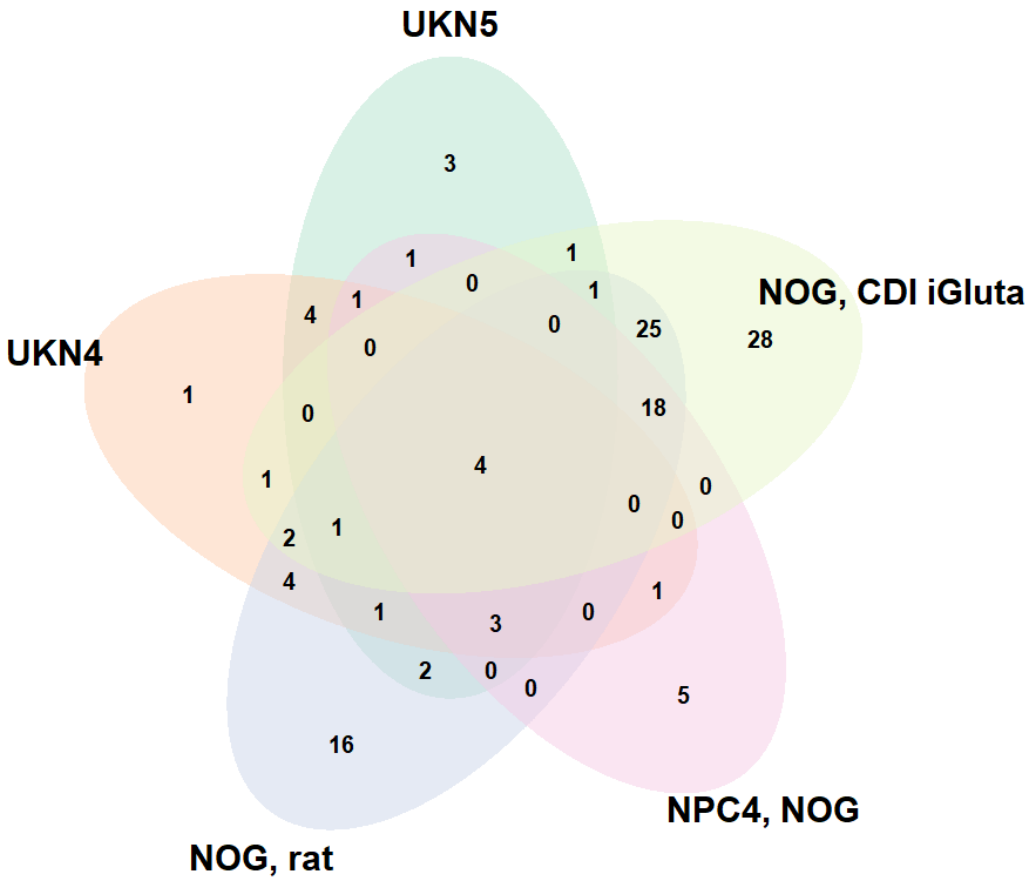


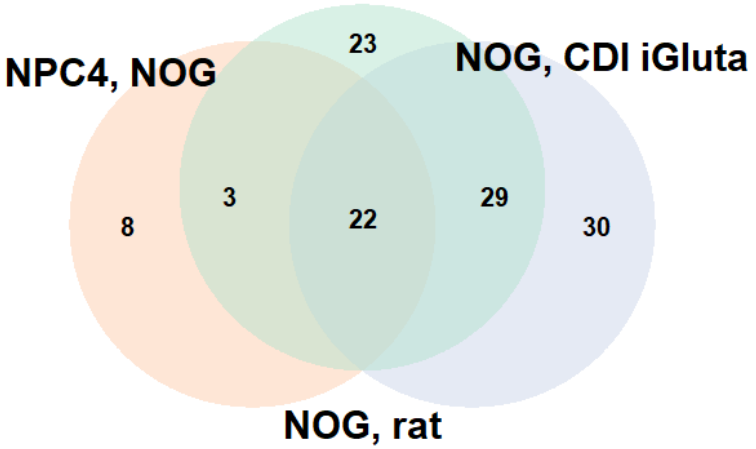
202 chemicals tested across all assays, including single-concentration and multi-concentration data

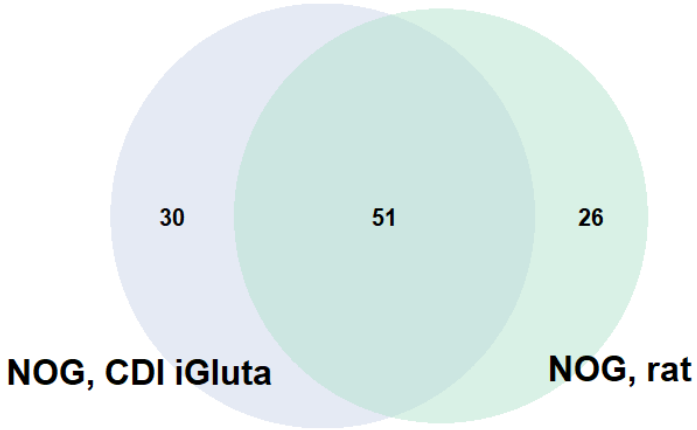
- Understand and explain differences in data from assays that measure the same developmental process.

Data Landscape: Hitcall agreement between assays measuring the same neurodevelopmental process

Neurite outgrowth

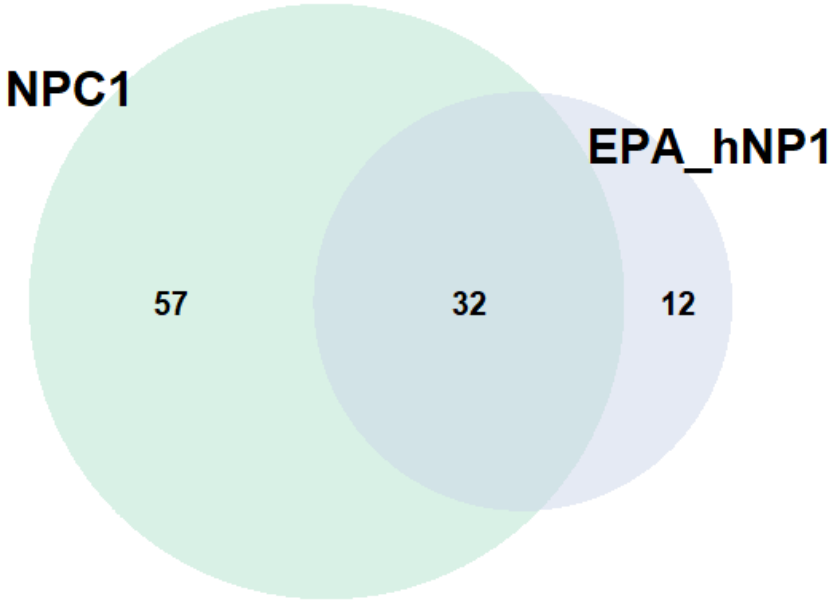




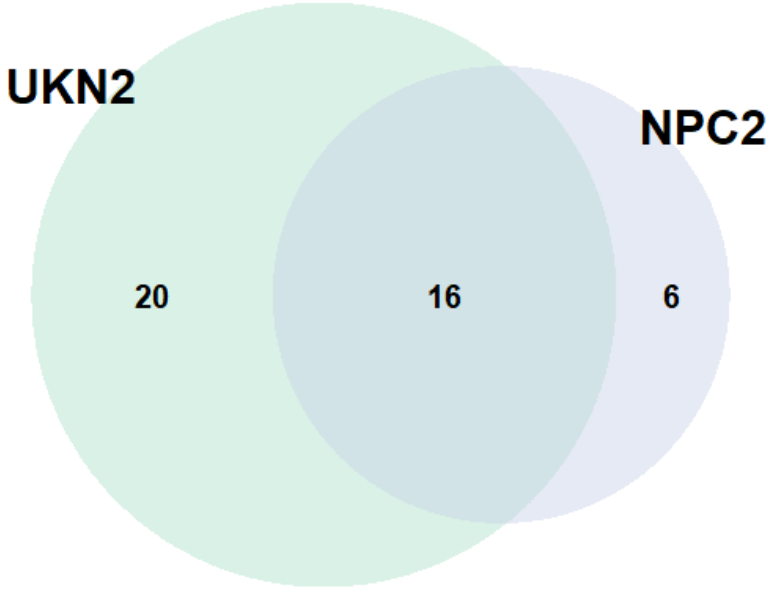


Data Landscape:
Hitcall agreement between assays measuring the same neurodevelopmental process

Proliferation



Migration (neuron)



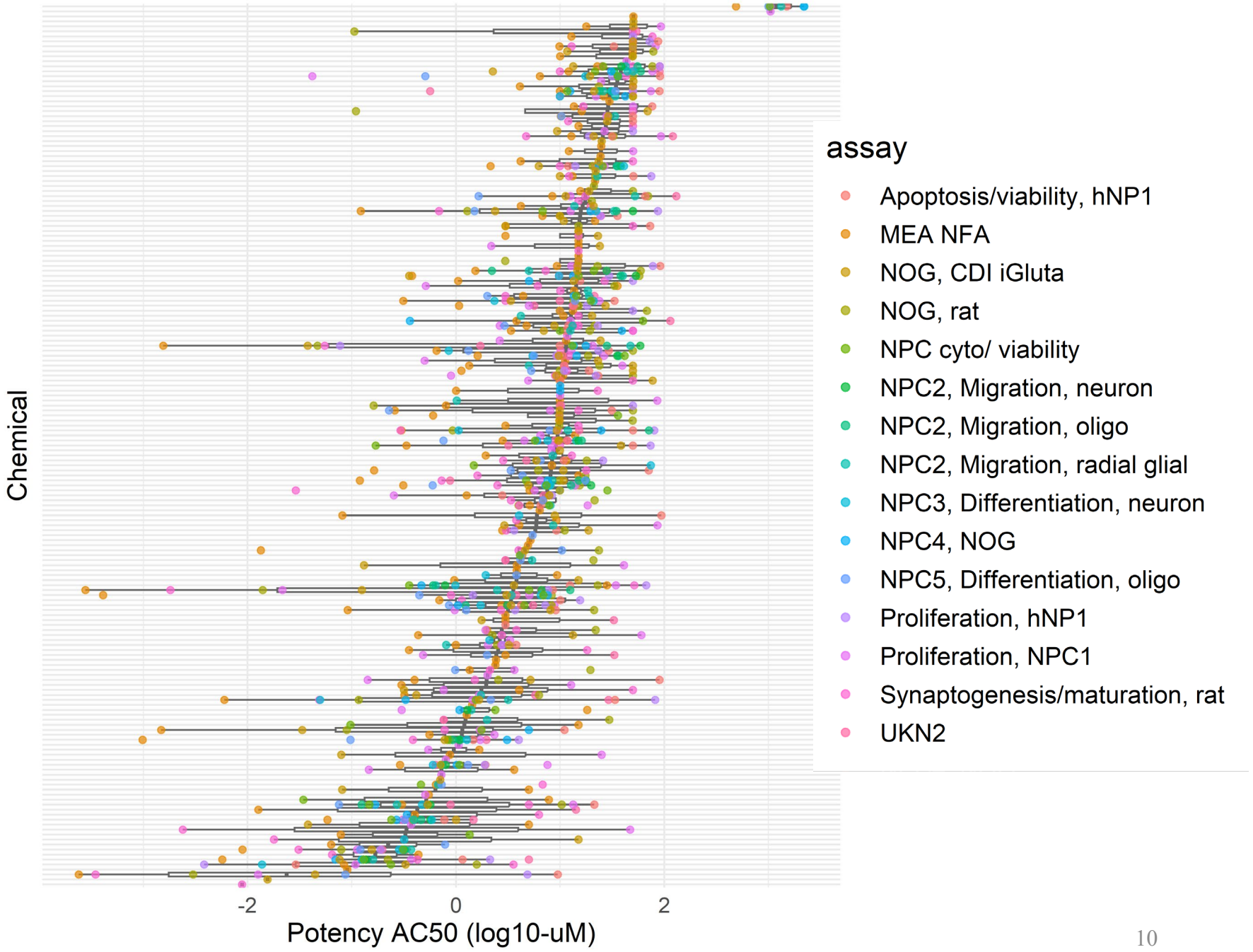
- Understand the sensitivity of some methods in the battery and advise for which purpose they should be used or not.

Data Landscape: Potency distribution across assays

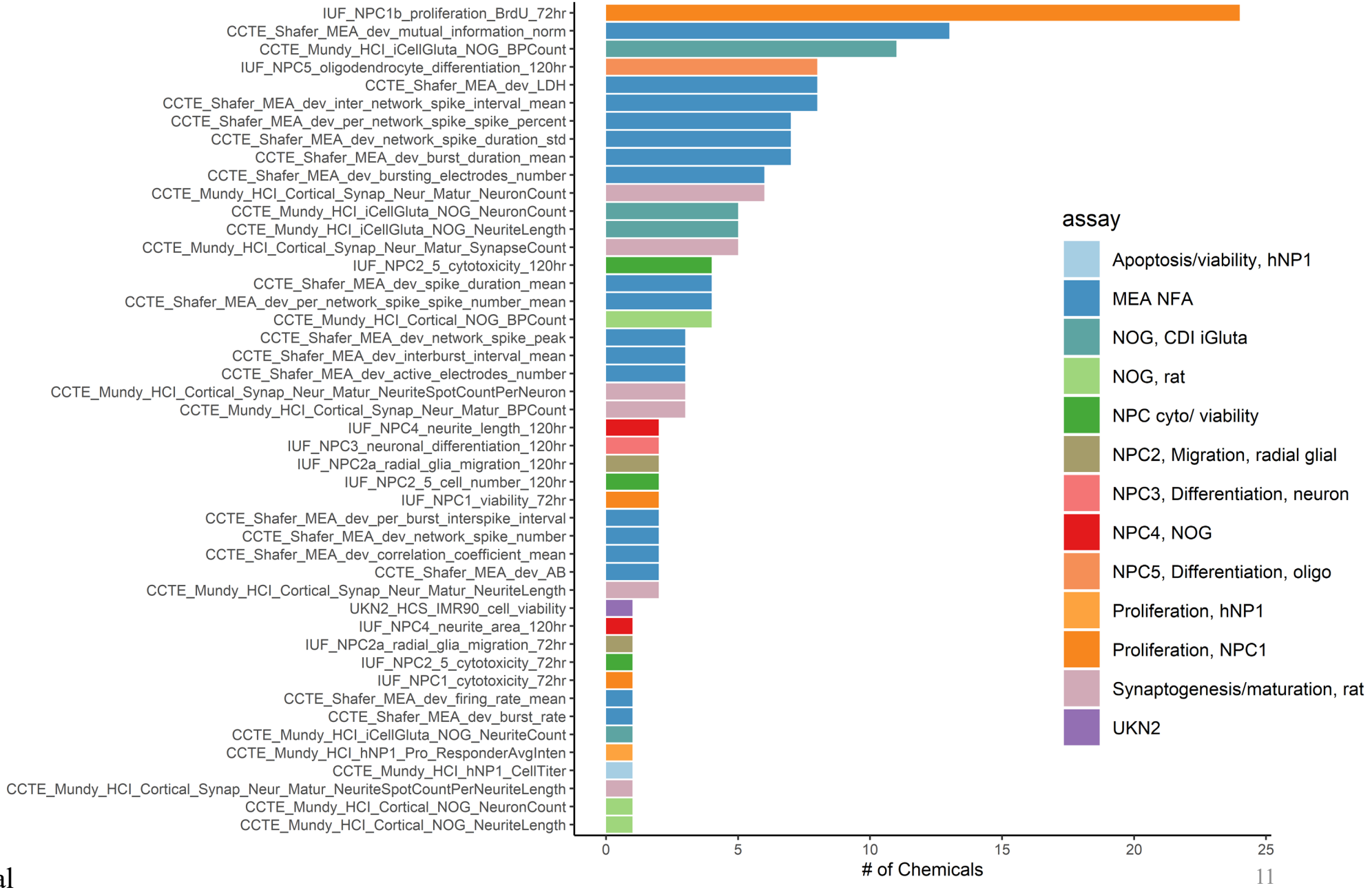
Mean potency range across 202 chemicals = 1.60 log10-uM (± 1.03 SD)

*Data points= minimum AC50 by assay

AC50= concentration at 50% maximal activity

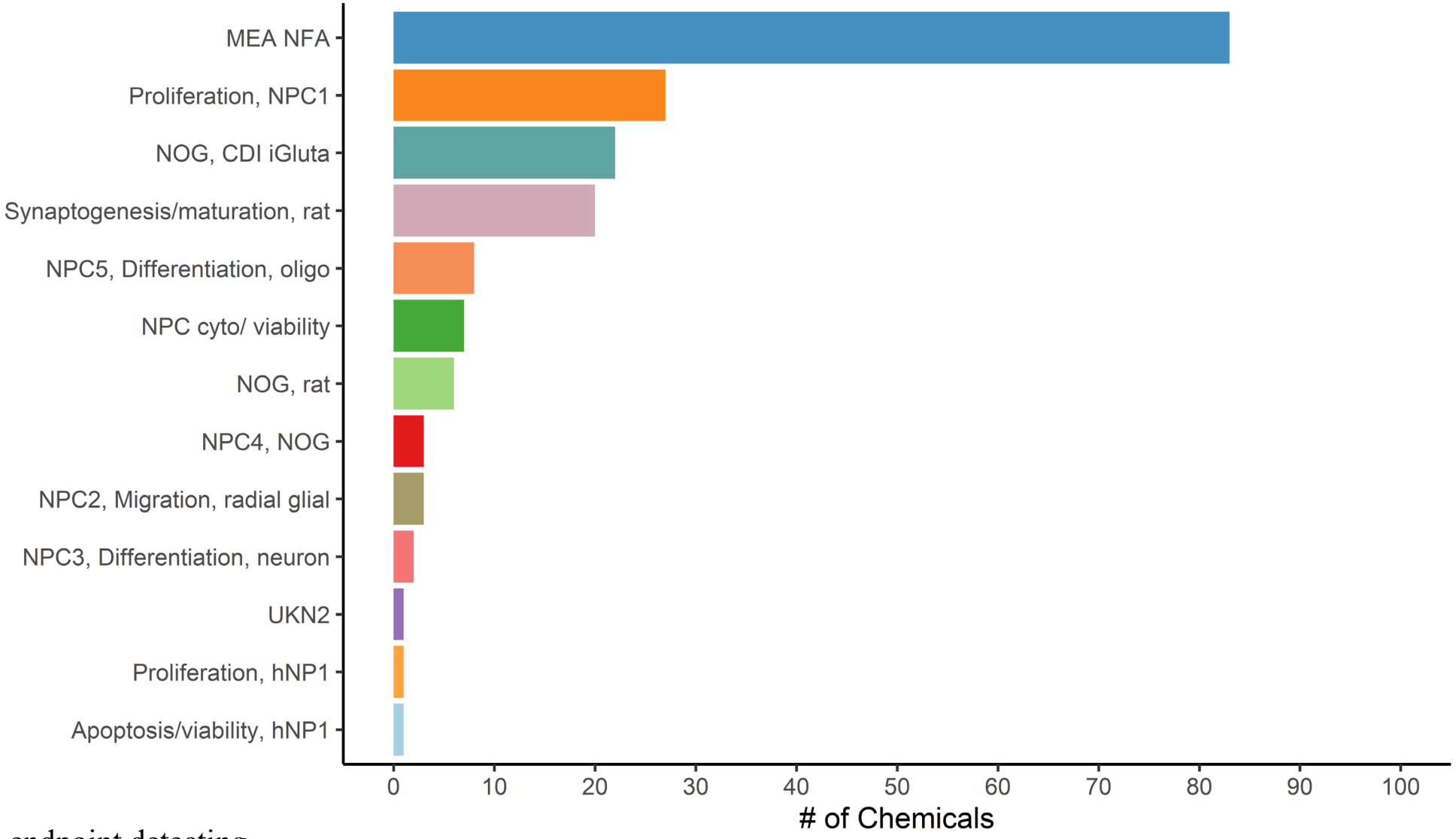


Data Landscape: Most sensitive endpoint



*Endpoint detecting the minimum AC50 by chemical

Data Landscape: Most sensitive assay

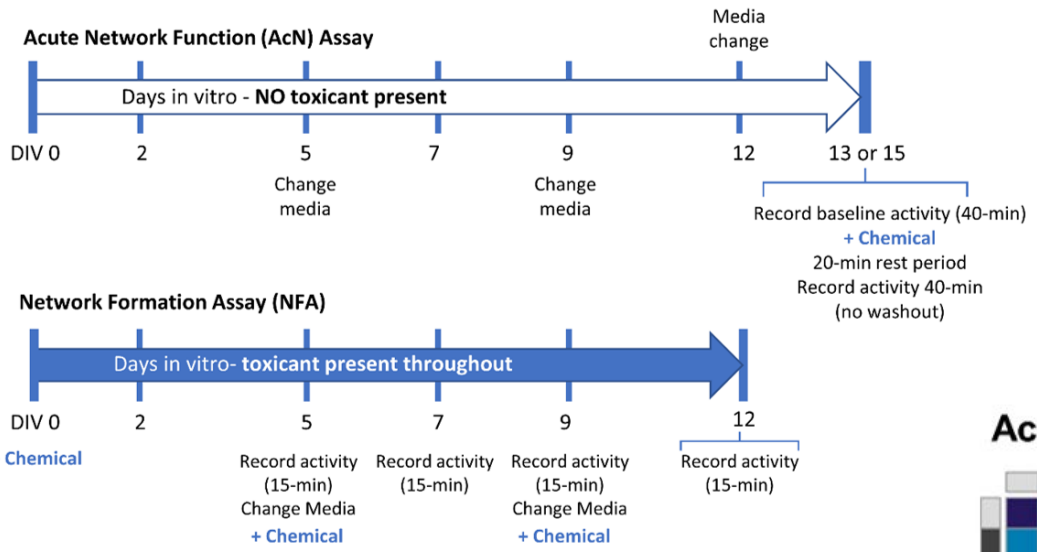


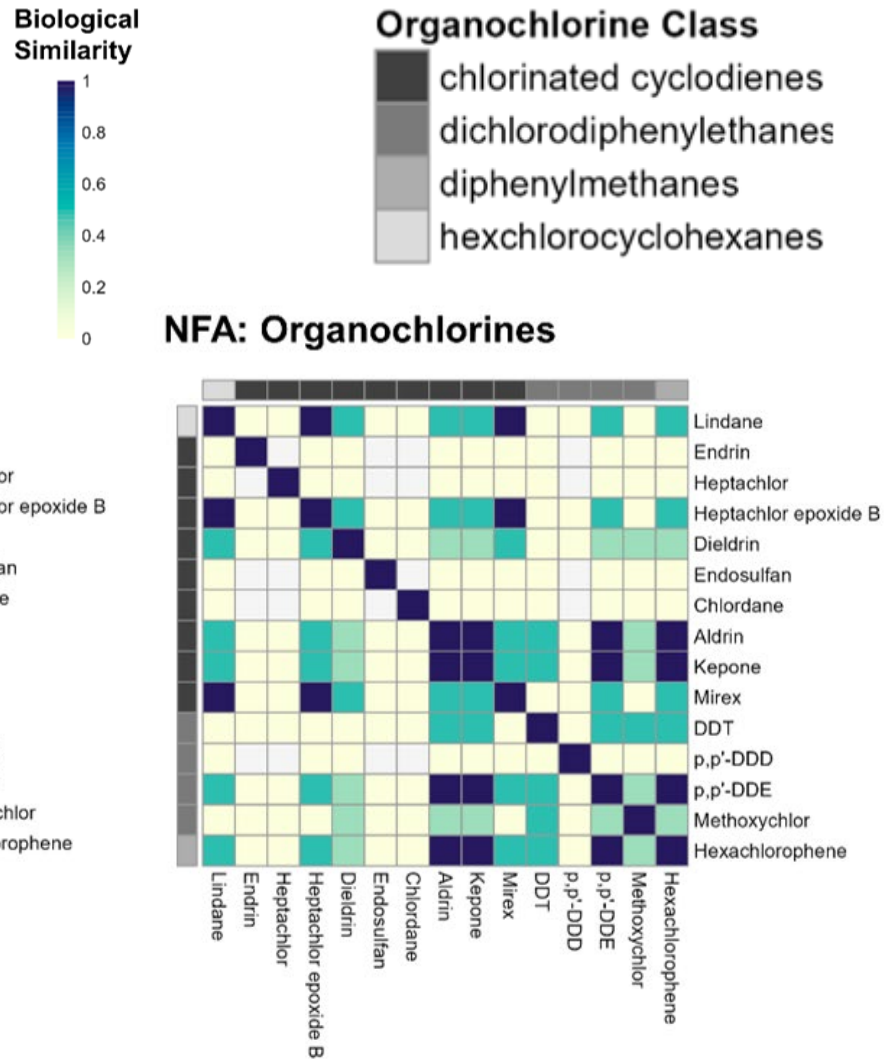
*Assay representing the endpoint detecting the minimum AC50 by chemical

Sharing methodologies to analyse data and ways to come up with one interpretation.

- Use machine learning to solve the problem of integrating data from the battery.

What can we learn about the biological interpretation of the MEA network formation assay (NFA) by comparing to the MEA acute exposure assay?

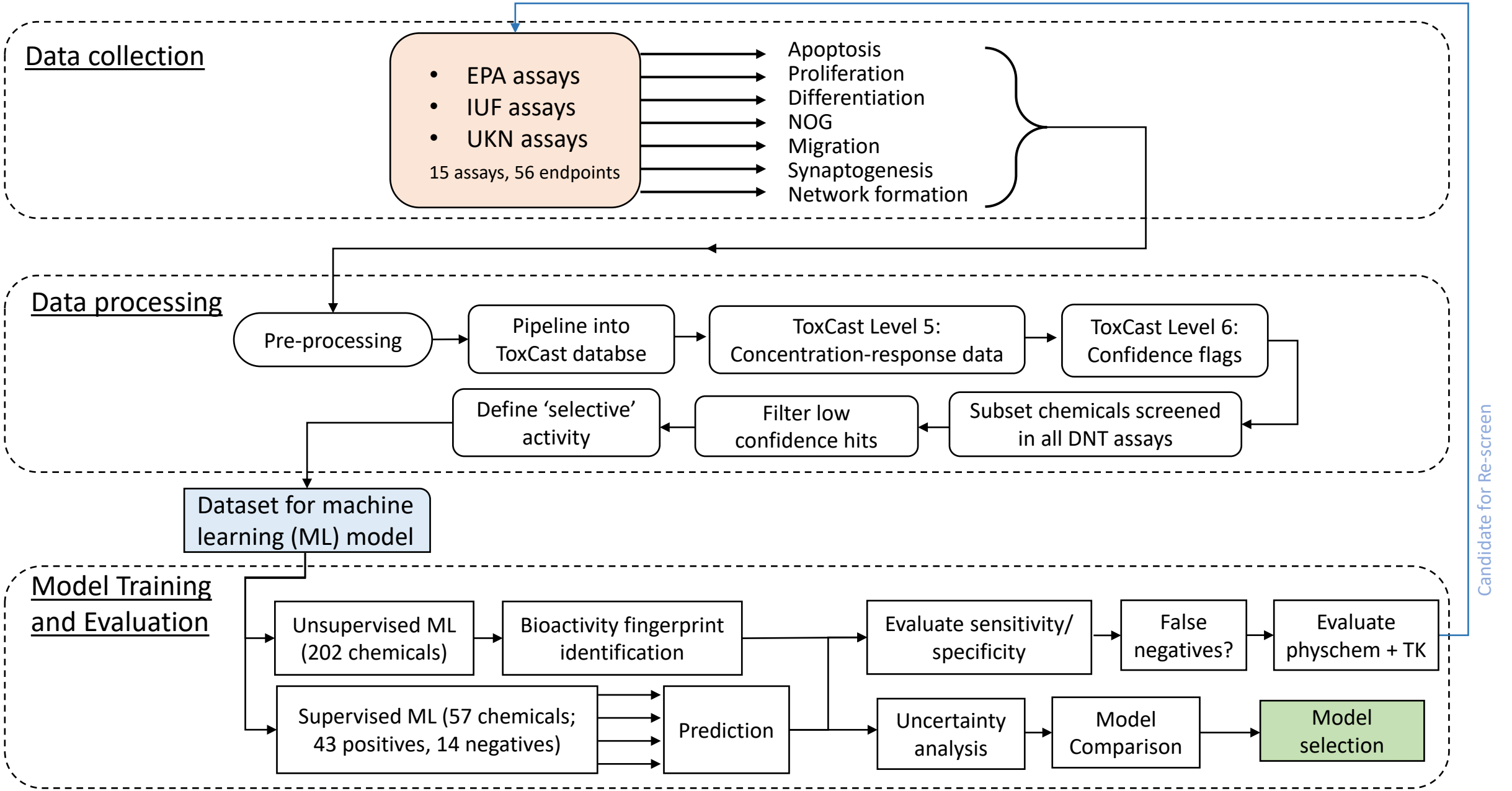




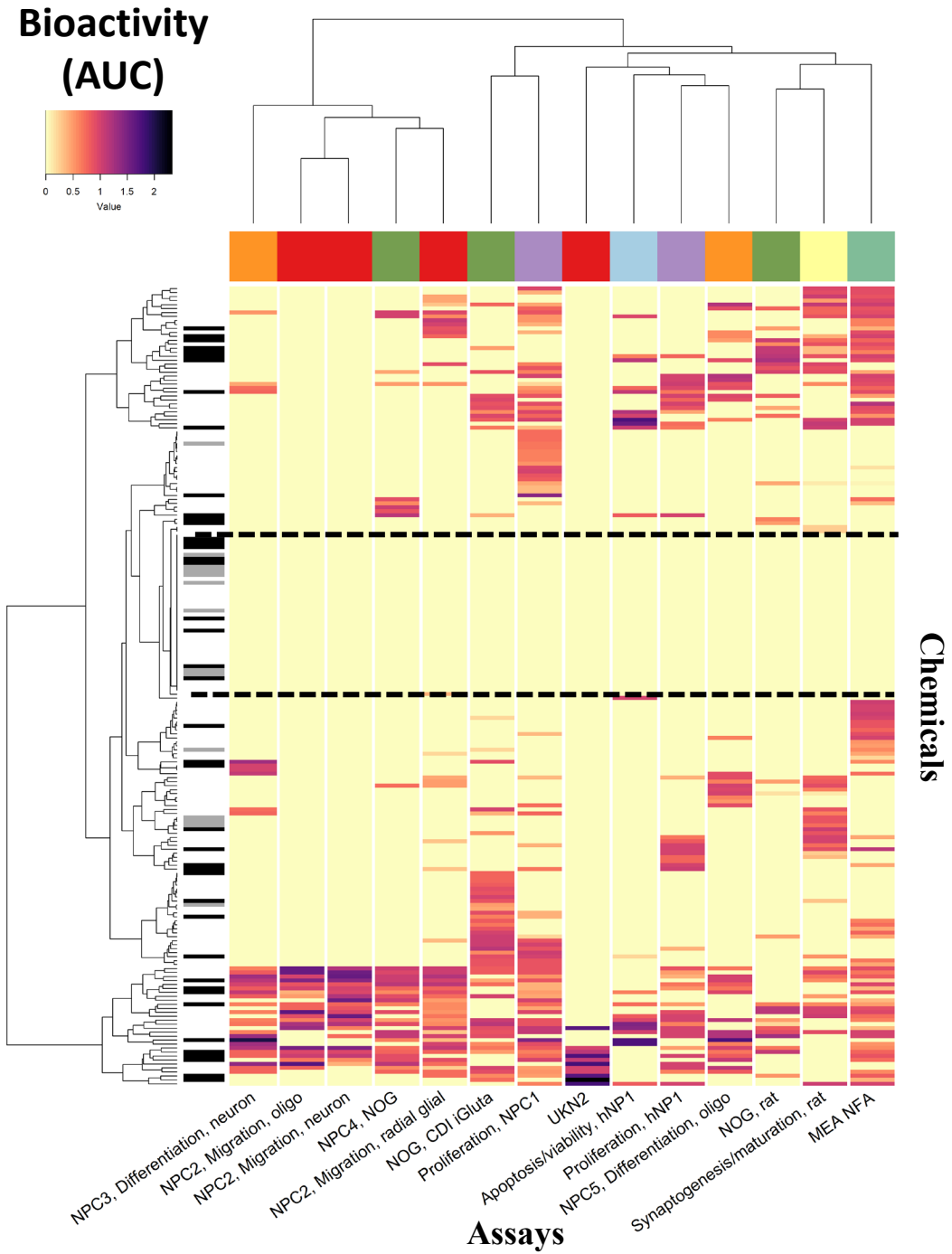
Biological similarity in the MEA NFA does not appear to be driven by a known mechanism of action, while bioactivity in the MEA acute does appear to be associated with mechanism of action.

Martin et al., (under revision)

Framework for AI-driven model for DNT



Model evaluation: unsupervised ML



Neurodevelopmental Process

- Apoptosis
- Network function
- NOG
- Migration
- Differentiation
- Proliferation
- Synaptogenesis

DNT *in vivo* evaluation chemical

- Positive
- Negative

- Clustering of assays appears to be driven by cell model rather than neurodevelopmental process.

AUC: area under the curve below the threshold of cytotoxicity, scaled sum by assay

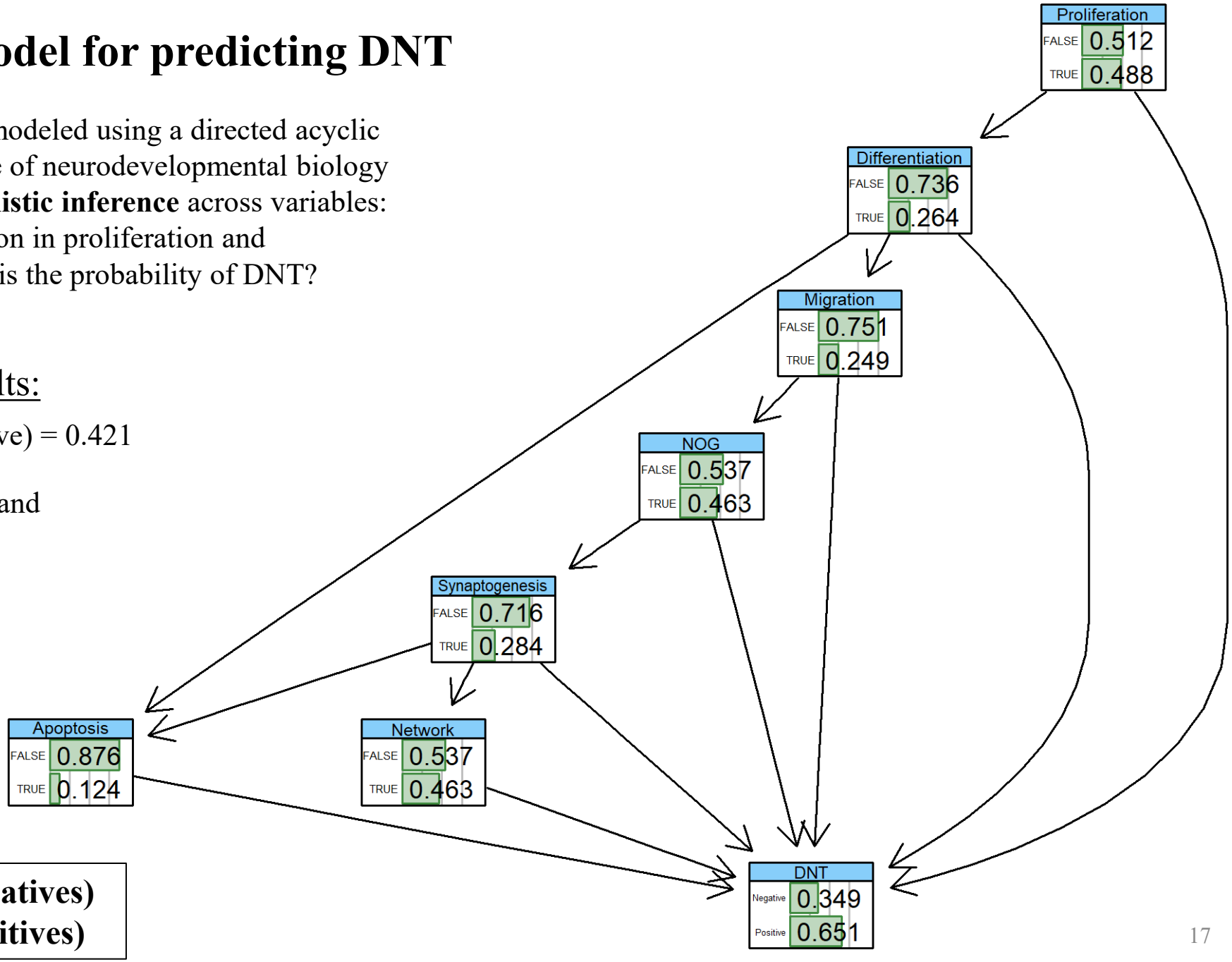
Bayesian network model for predicting DNT

- Conditional probabilities modeled using a directed acyclic graph and prior knowledge of neurodevelopmental biology
- Powerful tool for **probabilistic inference** across variables:
  - E.g. Given a disruption in proliferation and differentiation, what is the probability of DNT?

Preliminary inference results:

P(DNT Positive | Network is active) = 0.421

P(DNT Positive | Differentiation and Migration is active) = 0.304



- **81% sensitivity (8 false negatives)**
- **79% specificity (3 false positives)**

Acknowledgments

Data Contributions:
Tim Shafer lab
Ellen Fritsche lab
Marcel Leist lab

US EPA:
Biomolecular and Computational Toxicology Division

Computational Toxicology and Bioinformatics Branch

Division Direction: Sid Hunter
Branch Chief: Katie Paul Friedman

Contact Info:
Kelly Carstens†, PhD
U.S. Environmental Protection Agency
Research Triangle Park, NC

Email: carstens.kelly@epa.gov
Office: 919-541-3834

Supplemental

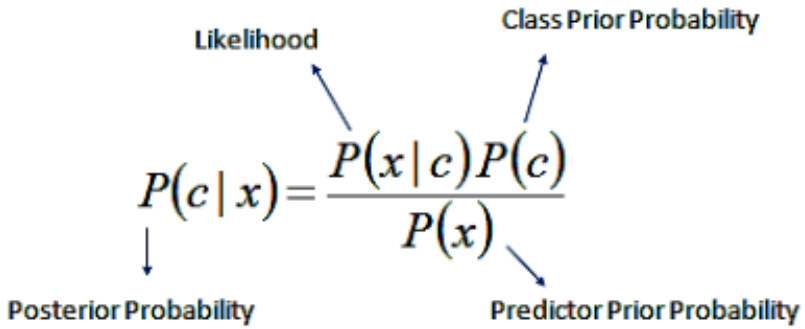
Points for discussion

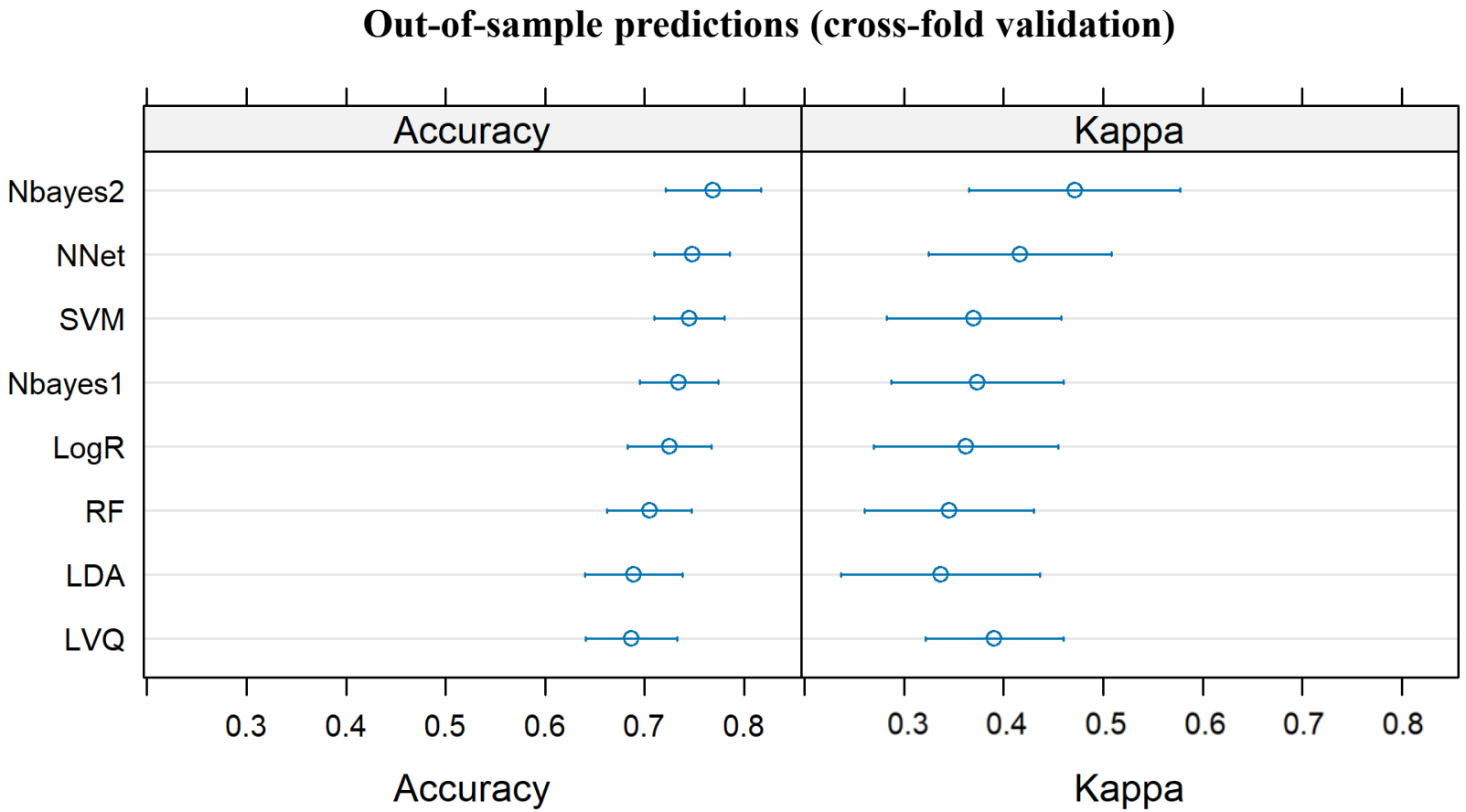
- Should we summarize assays measuring the same process for machine learning? (different cell types, species, etc.)
- Should we have different models based on fit for purpose? (prioritization versus biological interpretation)
- More chemical screening across assays are needed to evaluate the performance of machine learning techniques; what is the best approach for future chemical selection?
- Is classification of *in vivo* reference chemicals the best model outcome to inform the interpretation of the *in vitro* assays?
- What additional information would improve a machine learning model predict *in vivo* outcome? In vitro disposition data, toxicokinetic data, domain of applicability (e.g. target receptor present?), *in vivo* outcome specificity (motor or behavior)
- How do we determine if the assays space is 'good enough' in the context of a machine learning model?

Model evaluation: Supervised ML

Best performing dataset: Hitcall by neurodevelopmental process category (NOG, migration, etc.)

Best performing model: Naïve bayes with Laplace smoothing (handles zero probability outcomes) and attribute-weighting (AWNB) (helps overfitting, small datasets).





Confidence Level: 0.95

- **72% sensitivity (12 false negatives)**
- **93% specificity (1 false positives)**

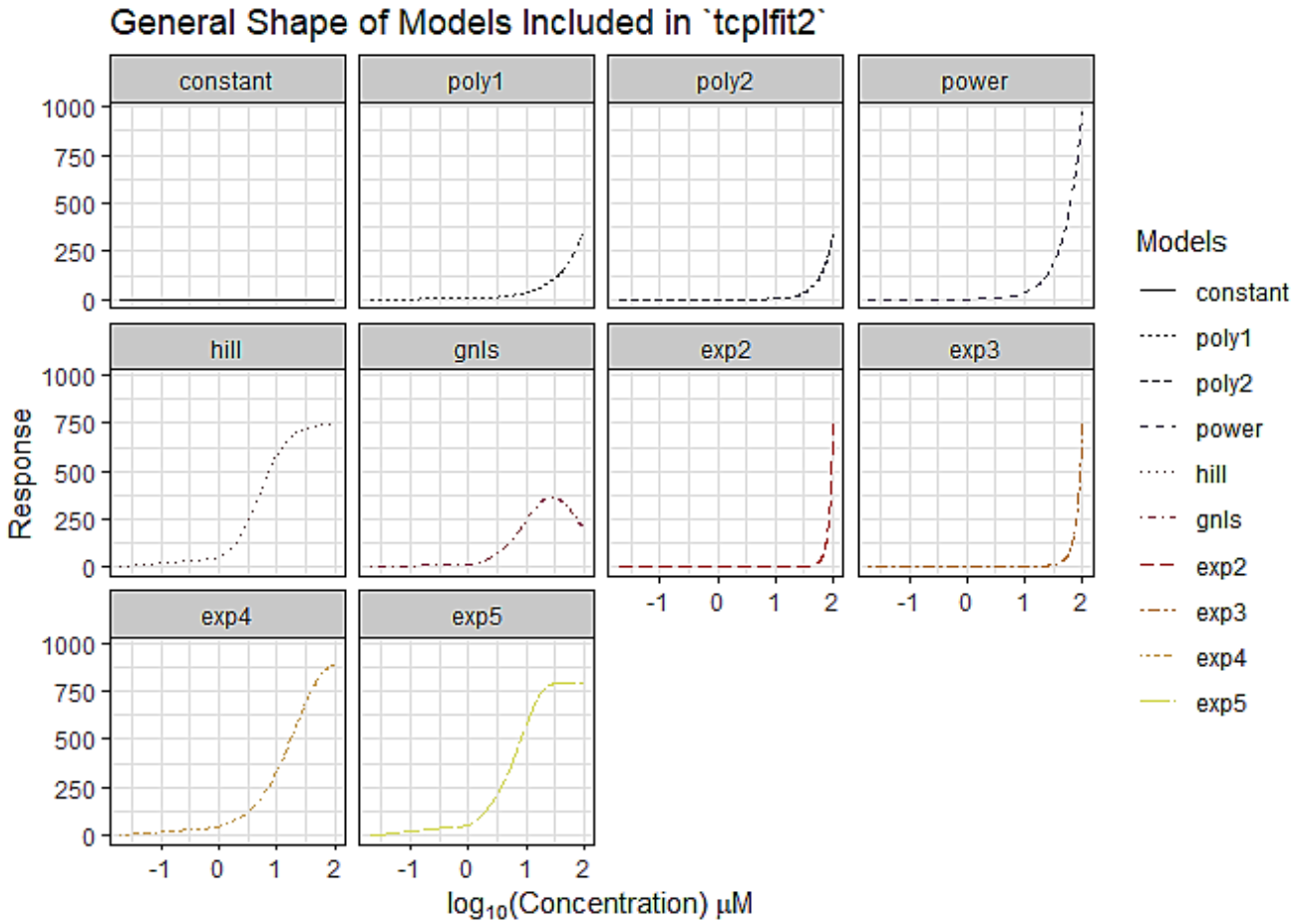
*Kappa: interrater reliability (observed vs expected accuracy)

ToxCast Pipeline data processing

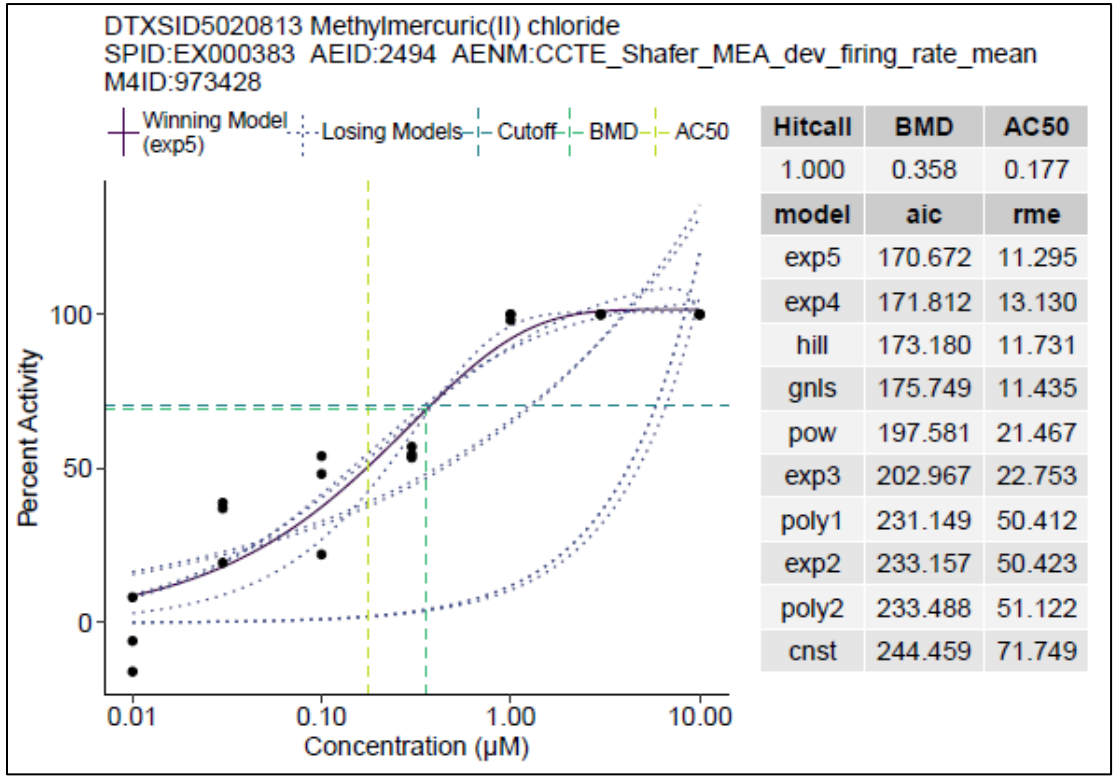
https://github.com/USEPA/CompTox-ToxCast-tcpl.git
R package: 'tcpl' v3.0

GitHub

Level 4: Model fitting



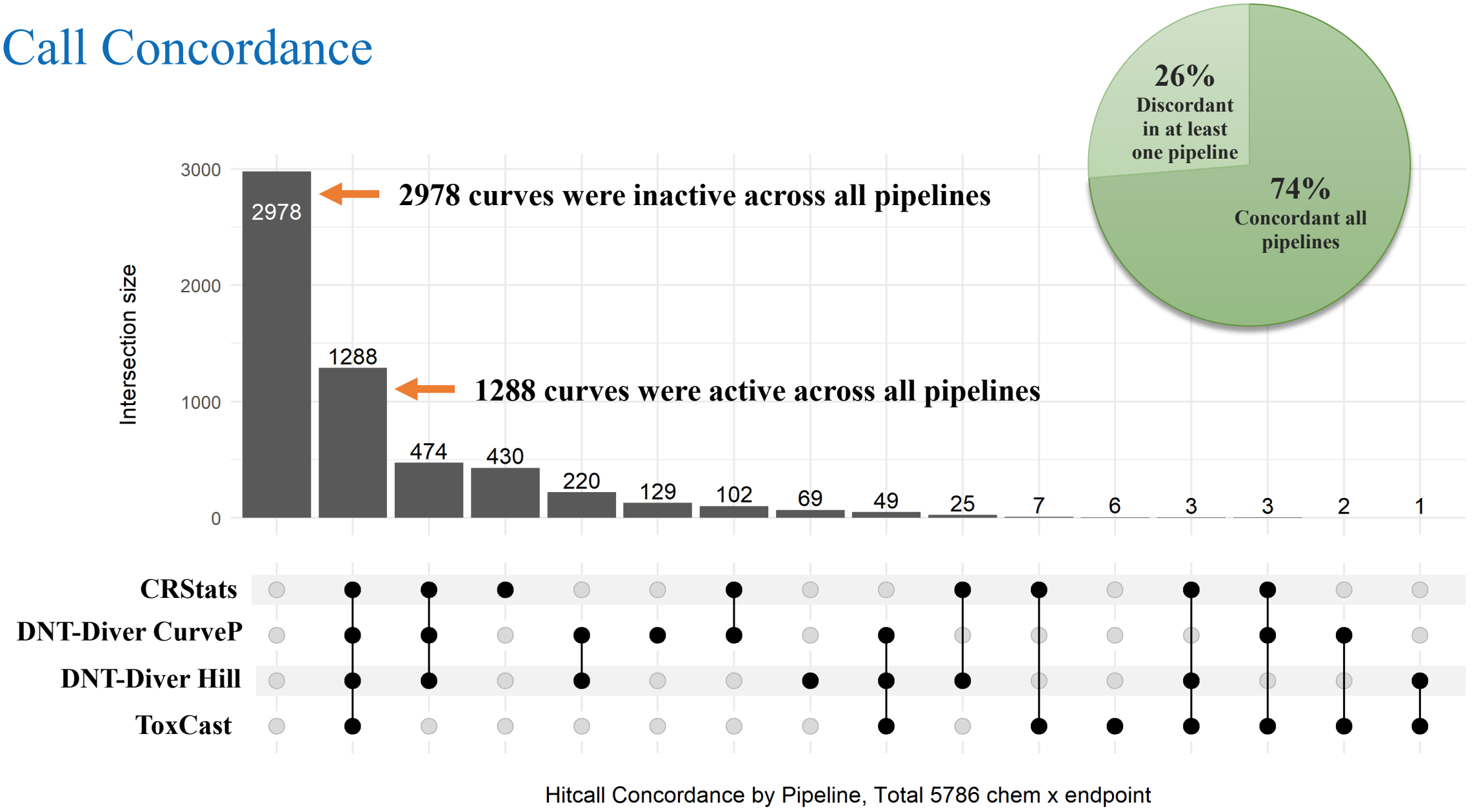
Level 5: Model Selection and hit calling



Hitcall	BMD	AC50
1.000	0.358	0.177
model	**aic**	**rme**
exp5	170.672	11.295
exp4	171.812	13.130
hill	173.180	11.731
gnls	175.749	11.435
pow	197.581	21.467
exp3	202.967	22.753
poly1	231.149	50.412
exp2	233.157	50.423
poly2	233.488	51.122
cnst	244.459	71.749

https://cran.r-project.org/web/packages/tcpl/vignettes/Data_processing.html

Hit Call Concordance



Compare two classification models for ‘selective’ activity

Method 1: CRStats Decision Tree

Utilizes confidence intervals between the cyto BMC and the DNT endpoint BMC to establish selectivity

(Keßel et al. ALTEX 2023)

Method 2: Selectivity Score

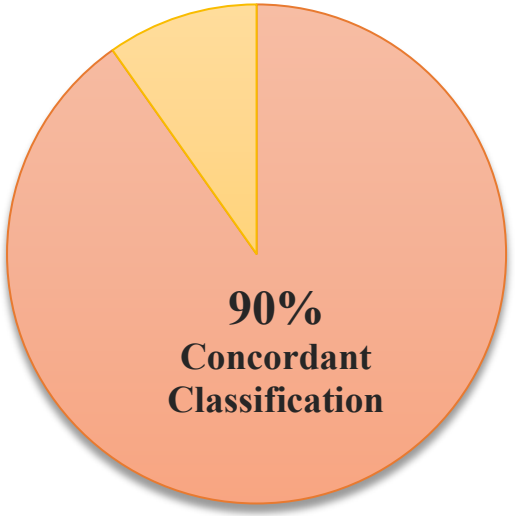
Score > 0.3 is selective

$Score = cyto\ POD\ (log10\ \mu M) - DNT\ endpoint\ (log10\ \mu M)$

(https://www.regulations.gov/document/EPA-HQ-OPP-2020-0263-0054)

Chemical Classification based on ‘selectivity’ analyses

		Method 2		
		Selective	Non-selective	Inactive
Method 1	Selective	76	5	0
	Non-selective	7	40	0
	Inactive	0	2	13



Of the 116 chemicals identified as active in both ‘selectivity’ methods, 12 chemicals did not agree on the classification as a ‘selective’ chemical for DNT.

Concentration-response modeling pipelines for DNT

- **ToxCast Pipeline**: US EPA
- **CRStats**: Leibniz Research Institute for Environmental Medicine, University of Konstanz
- **DNT-DIVER**: National Institute of Environmental Health Science (NIEHS) Division of Translational Toxicology (DTT)
- **PROAST**: RIVM National Institute for Public Health and the Environment, European Food Safety Authority

Acknowledgments

Jui-hua Hsieh[1]
Arif Donmez[2]
Martin Scholze[3]
Kristina Bartmann[2]
Ellen Fritsche[2]

EPA ToxCast Team:
Katie Paul Friedman[†]
Madison Feshuk[†]
Jason Brown[†]
Sarah Davidson[†]

[†]Center for Computational Toxicology and Exposure, ORD, US EP
1. National Institute of Environmental Health Sciences
2. Leibniz Research Institute for Environmental Medicine
3. Brunel University London

JOURNAL ARTICLE

tcpl: the ToxCast pipeline for high-throughput screening data FREE

Dayne L Filer, Parth Kothiya, R Woodrow Setzer, Richard S Judson, Matthew T Martin

Bioinformatics, Volume 33, Issue 4, February 2017, Pages 618–620,

Developmental **N**euro**T**oxicity **D**ata **I**ntegration and **V**isualization **E**nabling **R**esource (**DNT-DIVER**)

RIVM Committed to *health and sustainability*

Topics | About RIVM | Publications | International | Contact | Agenda

Home › PROAST

PROAST

GitHub ArifDoenmez / CRStats Public

Most active endpoints based on hit call