High-throughput H295R steroidogenesis assay: utility as an alternative and a statistical approach to characterize effects on steroidogenesis

Derik E. Haggard

ORISE Postdoctoral Fellow

National Center for Computational Toxicology

Computational Toxicology Communities of Practice

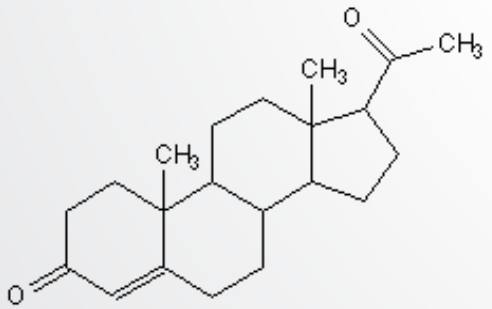
Dec. 14th, 2017

Outline

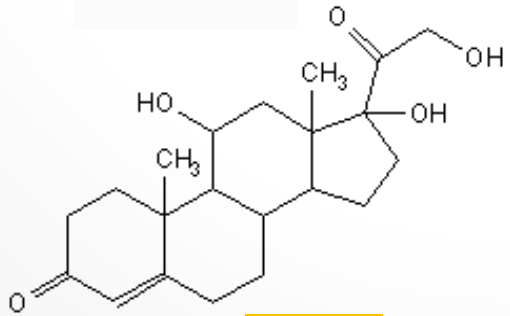
- Background
- Objectives
- Assay Background
- Methods and Results
  1. Evaluation of the HT-H295R assay
  2. Development of a quantitative prioritization metric for the HT-H295R assay data
- Summary and Conclusions

Steroid Hormone Biosynthesis & Metabolism

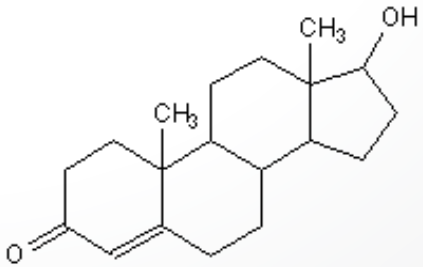
- Proper steroidogenesis is essential:
  - *In utero* for fetal development
  - In adults for reproductive function
  - Disruption can result in congenital adrenal hyperplasia, sterility, prenatal virilization, salt wasting, etc.
- >90% of steroidogenesis occurs in the gonads
  - Leydig cells (males) or follicular cells (females)
- Adrenal gland (corticosteroids)



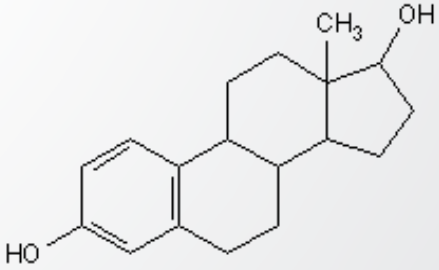
Progesterone



Cortisol



Testosterone

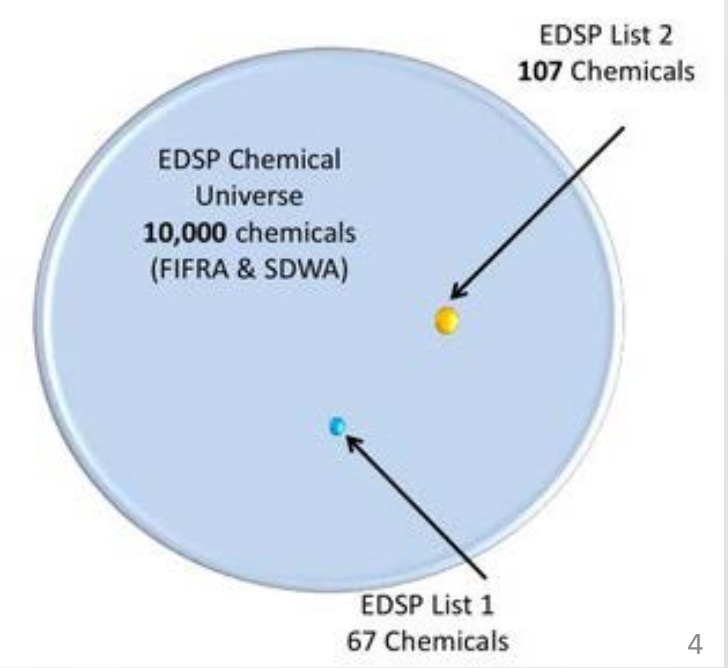


Estradiol

https://www.pharmacorama.com/en/Sections/Androgen_steroid_hormones.php

US EPA Endocrine Disruptor Screening Program (EDSP)

- EDSP mandated to screen chemicals for endocrine activity (estrogen, androgen, thyroid)
  - Initial tiered screen relied on low-throughput assays
- Modernization of EDSP (EDSP21) to use high-throughput and computational methods
  - Prioritize the universe of EDSP chemicals for endocrine bioactivity
- Altering hormone levels via disruption of biosynthesis or metabolism can also contribute to endocrine disruption
  - This is difficult to assay *in vitro*
- Current Tier 1 Assay:
  - OECD-validated **H295R steroidogenesis assay**



https://www.epa.gov/sites/production/files/2015-04/documents/edsp_dix_ord_communities_of_practice_04_23_15_f.pdf

Objective 1

- *EPA and OECD test guidelines for the H295R steroidogenesis assay to detect potential perturbation of estradiol (E2) and testosterone (T) synthesis are designed for low-throughput screening*

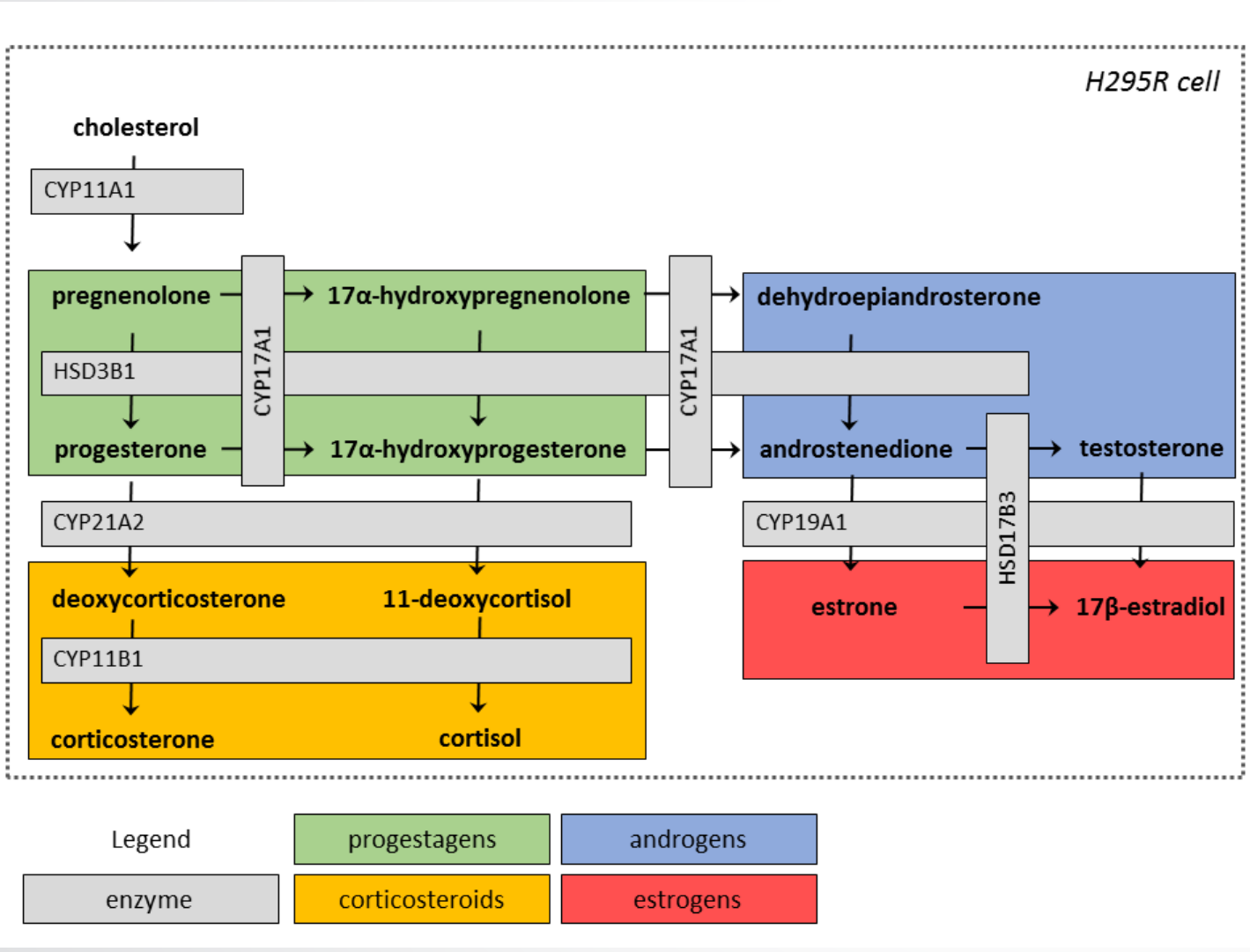
- **Objective 1:** Compare the recently developed high-throughput H295R assay (HT-H295R; refer to Karmaus *et al.,* 2016) to the OECD test guideline assay

Objective 2

- *The HT-H295R assay includes measurement of 13 hormones to represent the steroidogenesis pathway*

- **Objective 2:** Develop a summary measure that integrates these multi-dimensional data to quantify pathway perturbation and indicate relative priority for further screening and/or evaluation of chemicals for potential effects on steroidogenesis

High-throughput Steroidogenesis Assay in H295R (HT-H295R)



- Assay contract with Cyprotex (formerly CeeTox) and OpAns
- Adrenocortical carcinoma cell line
  - All of the major biosynthetic enzymes for steroidogenesis present
- Characteristics of undifferentiated fetal adrenal cells
  - Change steroid hormone output based on culturing conditions
- Measure production of up to 13 hormones/intermediates
  - HPLC-MS/MS

HT-H295R Assay Method

Concentration-Response

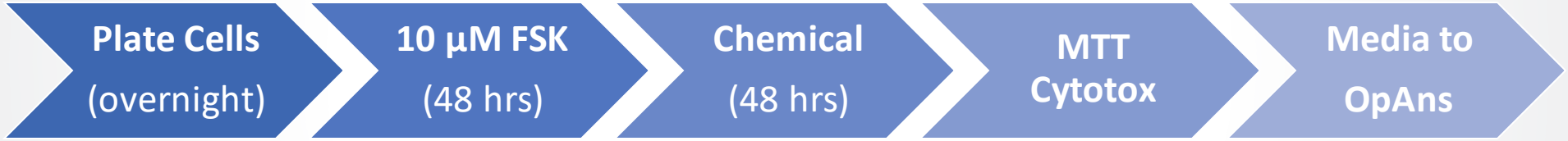


Plate Cells (overnight)	10 µM FSK (48 hrs)	Chemical (48 hrs)	MTT Cytotox	Media to OpAns
H295R cells seeded to ~50% confluency	pre-treatment: stimulate steroidogenesis	6 conc. chemical treatment	Cell viability (≥70%)	HPLC-MS/MS quantification of hormones

Objective 1: Evaluation of the HT-H295R Assay

Comparison of results to the reference chemicals used for the OECD inter-laboratory validation

Does the HT-H295R Assay Replicate Results of the OECD H295R Assay?

- Comparison to the reference chemicals used in the OECD inter-laboratory validation study (Hecker *et al*., 2011)

Major differences between assays:

Primary Difference	OECD H295R	HT-H295R
Number of chemicals (multiple concentrations)	28	656
Cell culture	24 hr. plating, then 48 hr. exposure (total = 72 hr.)	Overnight plating, 48 hr. **forskolin pre-stimulation**, 48 hr. exposure (total = 112 hr.)
Cell viability threshold	≥80%	≥70%
Number of steroids measured	2	13
Quantification method	ELISA or LC-MS	HPLC-MS/MS

HT-H295R Data Analyzed Using Methods from OECD Inter-laboratory Validation (Hecker et al. 2011)

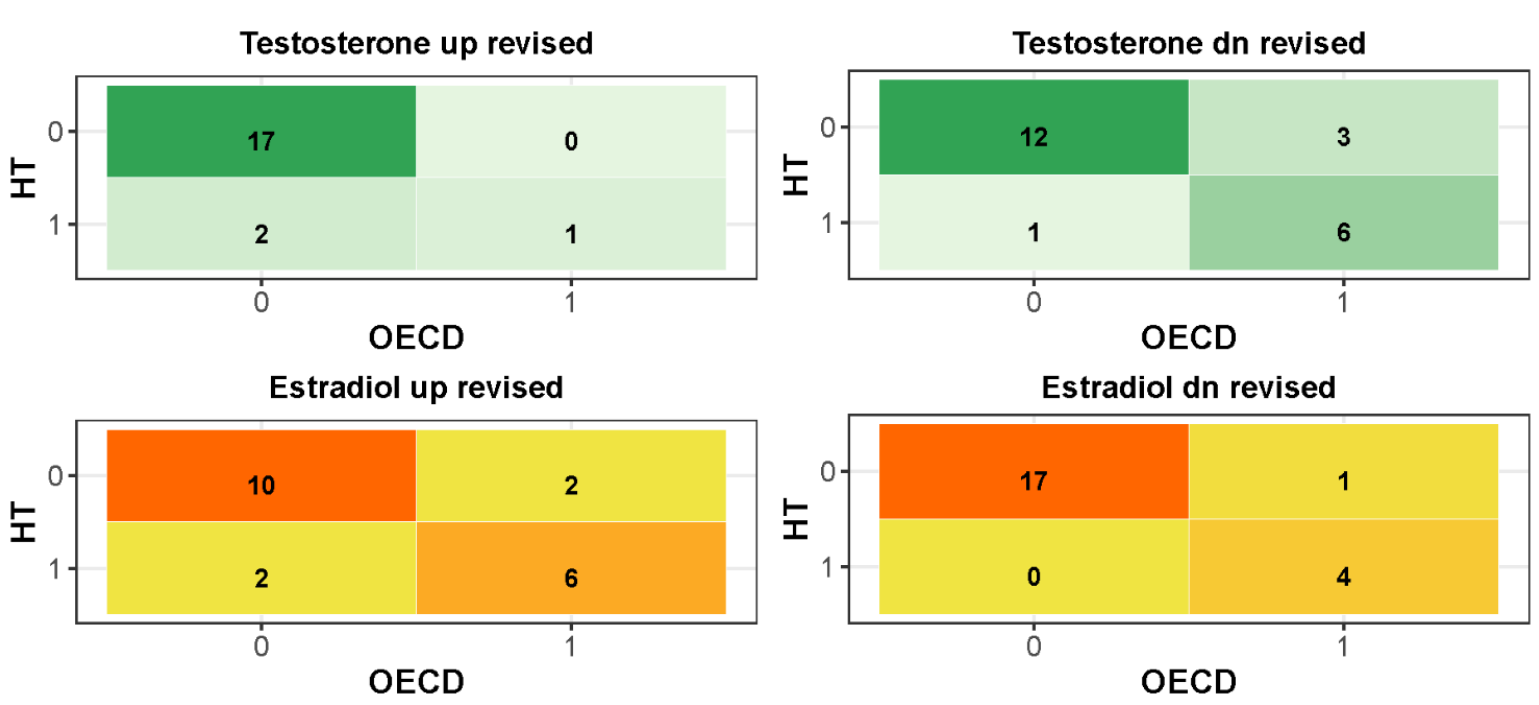
- ANOVA and Dunnett's with $\alpha = 0.05$
- DMSO control data from the same plate were used for the sample comparison
- Criteria for positive:
  - 2 consecutive concentrations had to produce results significantly different from control
  - Or, positive at the max concentration that maintained ≥ 70% cell viability
  - 1.5-fold change from DMSO control was applied

Constructing Confusion Matrices

- 10/12 core reference chemicals shared
  - Tested in 5 labs for the inter-laboratory validation
- 15/16 supplemental reference chemicals shared
  - Tested in 2 labs for the inter-laboratory validation
- OECD inter-laboratory results were equivocal and removed if: ≥ 2 of 5 labs failed to report a LOEC (core reference chemicals) or 1 of 2 labs failed to report a LOEC (supplemental reference chemicals)

Confusion Matrices Demonstrate Good Sensitivity, Specificity, and Accuracy for Reference Chemicals.

Effect	Revised Sensitivity	Revised Specificity	Revised Accuracy
Testosterone up	1.00	0.89	0.90
Testosterone dn	0.67	0.92	0.82
Estradiol up	0.75	0.83	0.80
Estradiol dn	0.80	1.00	0.95



Agreement Among Labs in the Inter-laboratory Validation

- For any effect on **testosterone**:
  - Average concordance among labs was 0.88, 0.91, and 0.90 for the 12 core reference chemicals only, the 16 supplemental reference chemicals only, and the entire set.

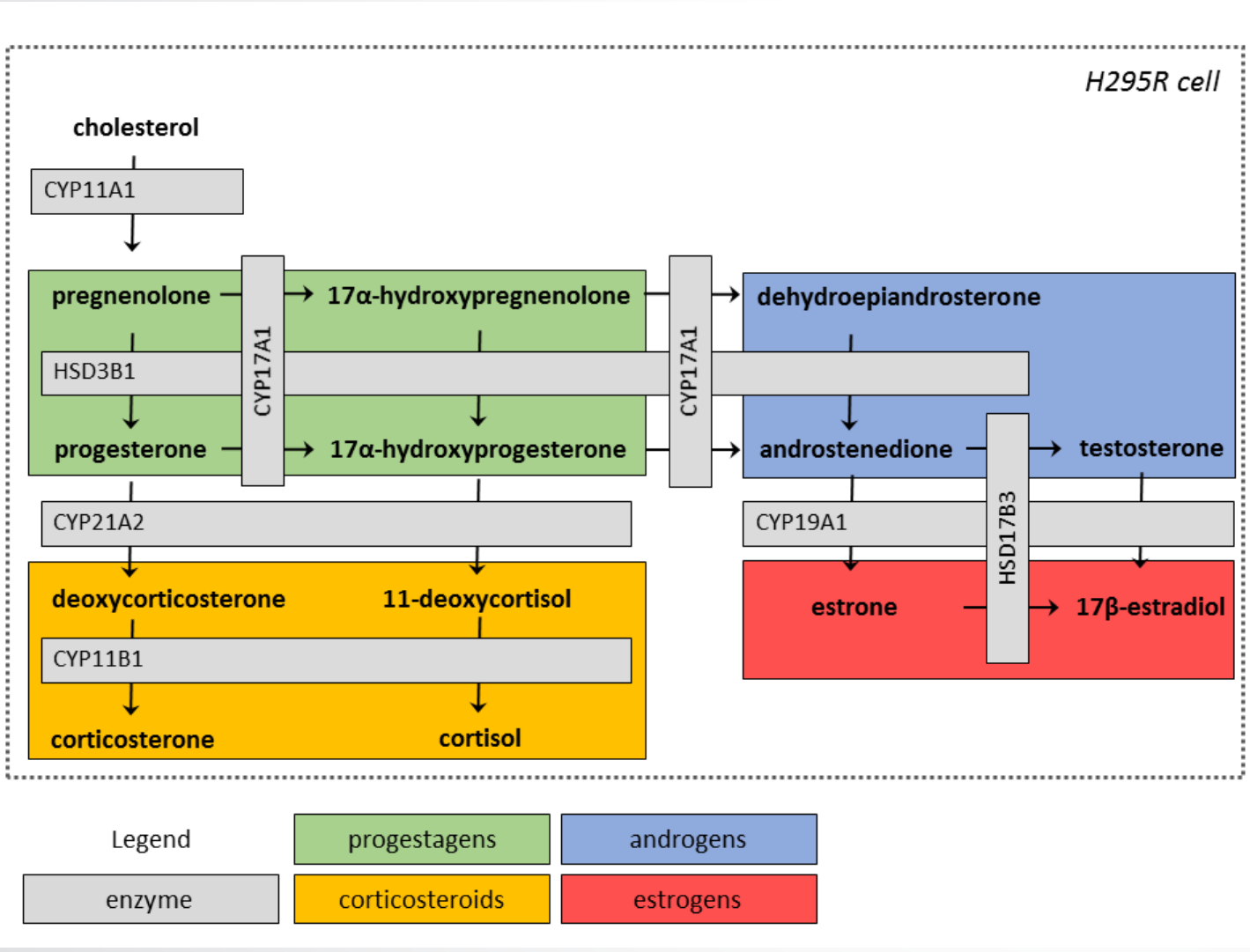
- For any effect on **estrogen**:
  - Average concordance among labs was 0.95, 0.84, and 0.89 for the 12 core reference chemicals only, the 16 supplemental reference chemicals only, and the entire set.

- Similar concordance between the HT-H295R and the OECD inter-laboratory validation

Objective 2: Development of a Quantitative Prioritization Metric for the HT-H295R Assay

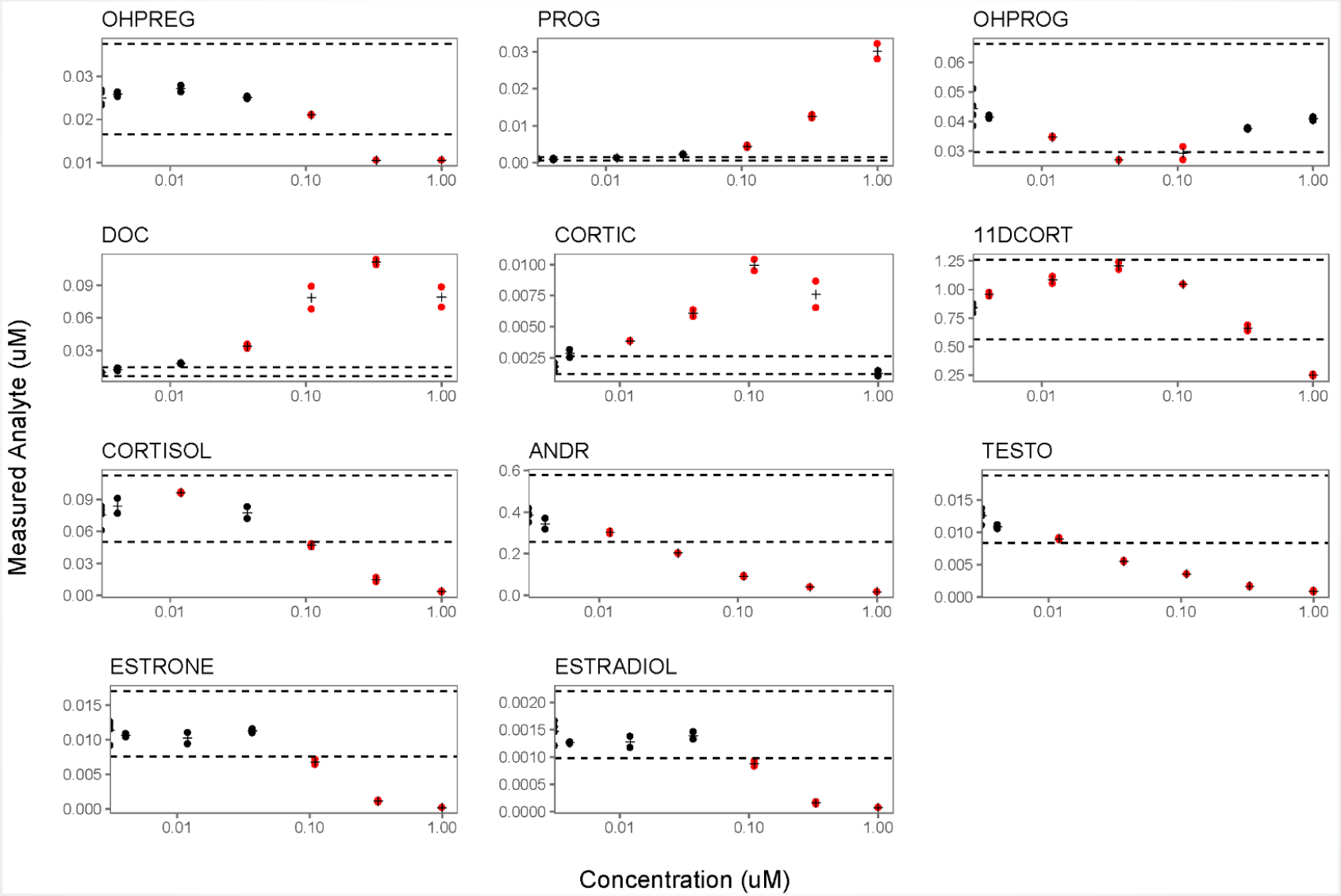
Simplifying an 11-dimensional problem to 1-dimension for prioritization

HT-H295R Data for Development of a Prioritization Metric



- 13 hormones measured in HT-H295R
- Pregnenolone and DHEA were often measured ≤ LLOQ (53.1% and 69.5% of all measurements) and were excluded

Example of the 11-dimensional Results for Prochloraz



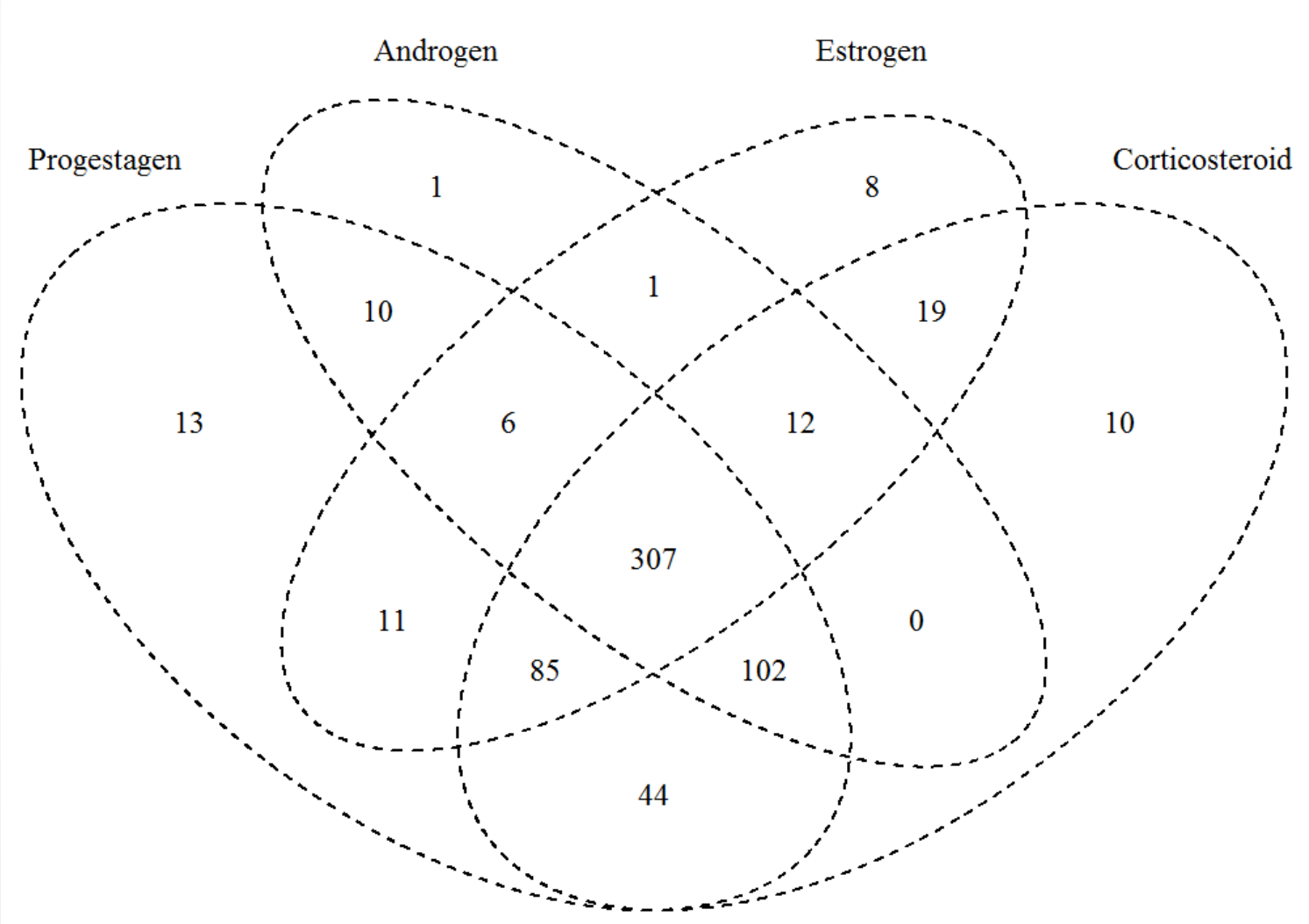
What Can We Learn from the Other Steroid Hormone Data Available in HT-H295R?

- Additional evidence for disruption of estrogen or androgen synthesis (e.g., estrone and androstenedione disruption)
- Putative mechanisms of steroidogenesis disruption
- Information about effects on other specific steroid hormone classes, namely the corticosteroids and progestagens

(Asser et al., 2014; Nielsen et al., 2012; Tinfo et al., 2011; Zhang et al., 2011)

Chemicals Screened in HT-H295R Had a Variety of Effects on Steroid Biosynthesis

629 chemical samples (out of 654 total) affected production ≥ 1 steroid hormone class.



Developing a Prioritization Metric for HT-H295R

Goal: Integrate HT-H295R data into a single value which estimates the overall magnitude of perturbation of steroidogenesis in H295R cells

Challenges:

- Multivariate dataset (11 hormone measures per chemical per concentration)
- Hormones are measured from the same experimental well
- Concentrations of steroid hormones and intermediates are often interdependent

A Simple Solution: Euclidean Distance

- Euclidean distance is a measure that can be used to estimate the distance between two points in multivariate space:

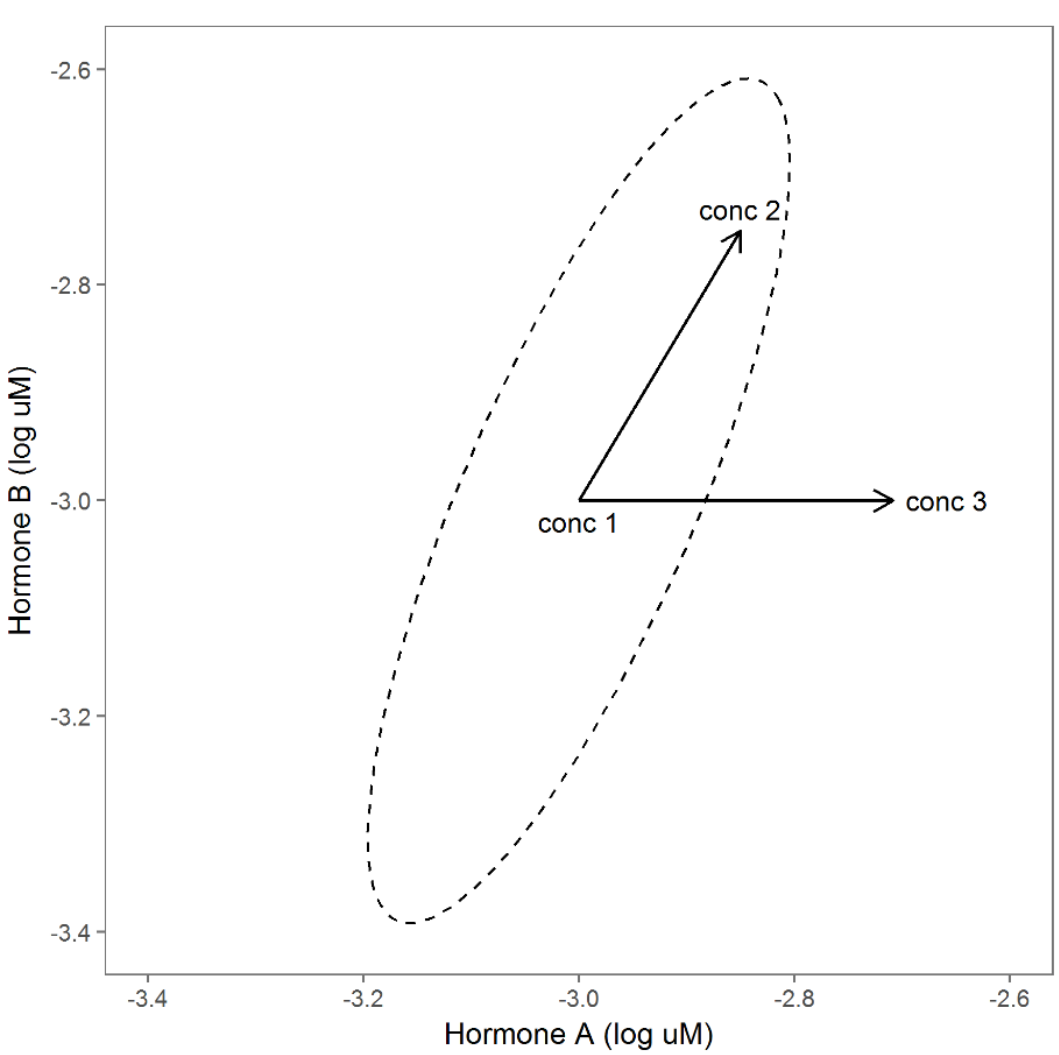
$$d = \sqrt{(y_c - y_1)^T (y_c - y_1)}$$

Where:

- y_c is the vector of natural log-transformed steroid hormone concentrations at the c^{th} concentration
- y_1 is the vector of natural log-transformed steroid hormone concentrations for the DMSO control

http://mccormickml.com/2014/07/21/mahalanobis-distance/

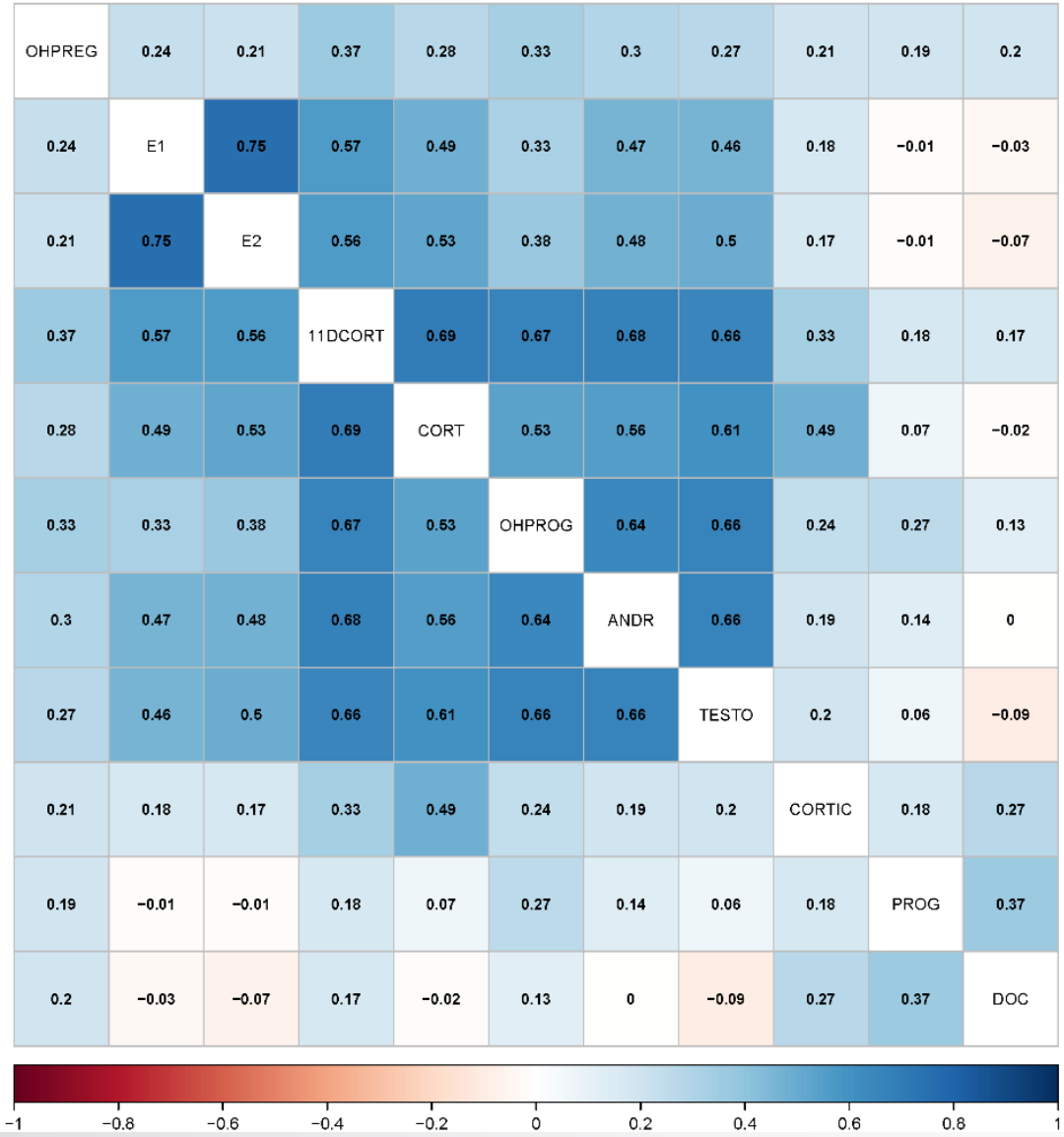
Limitations of Euclidean Distance



Conceptual Example:

- Hormone A and B show positive covariance
- conc 2 and conc 3 have the same Euclidean distance from conc 1
- Even though conc 3 is more standard deviations, i.e. a more ‘extreme’ distance from conc 1 than conc 2

The Residuals for Some Steroid Hormones in HT-H295R are Correlated

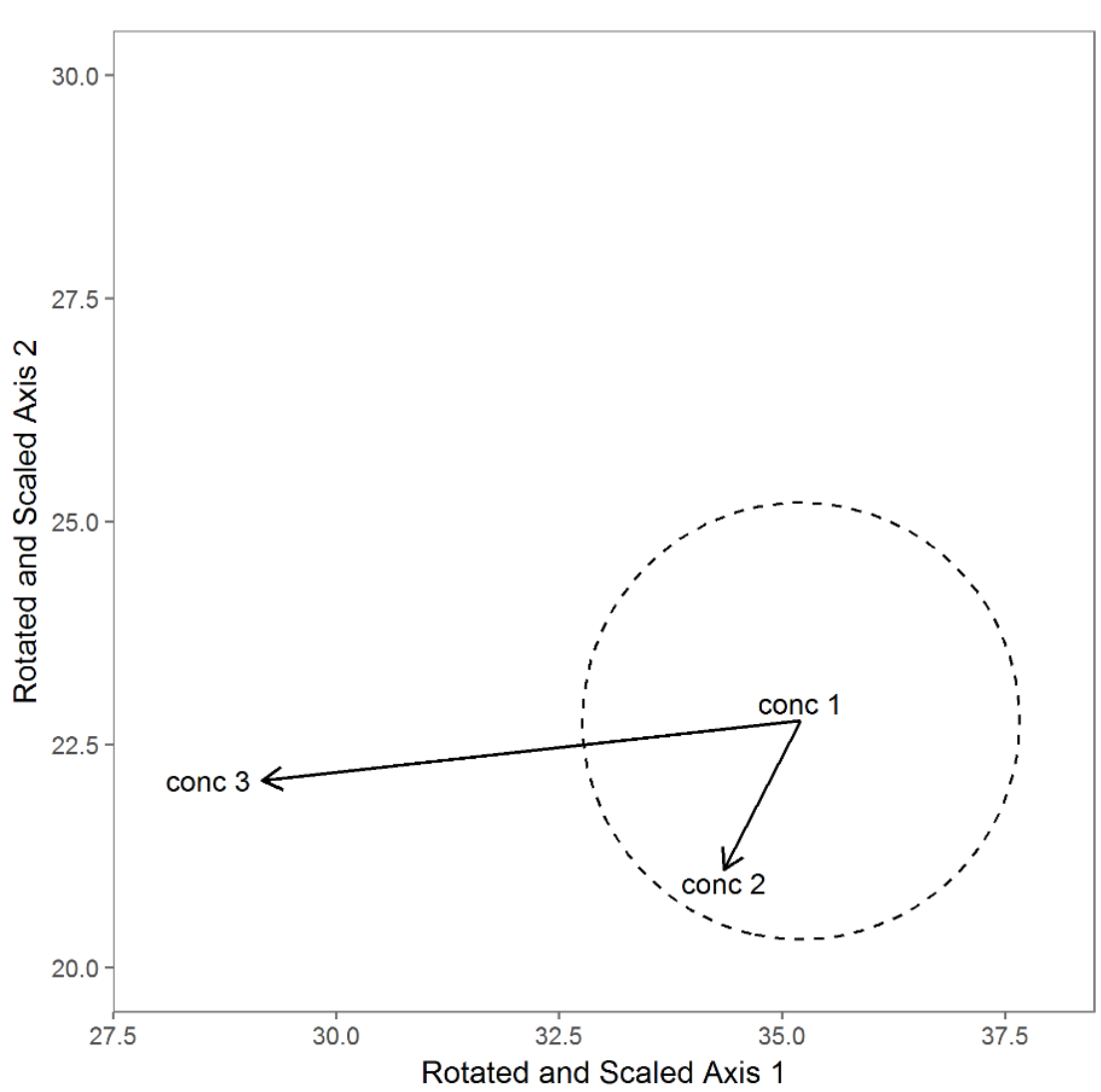


Highly correlated residuals:

- Estrone and E2 (Pearson's R = 0.75)
- Androstenedione and T (R = 0.66)
- Cortisol and 11-deoxycortisol (R = 0.69)

- Euclidean distance not appropriate for HT-H295R

Removing Residual Covariance Using the Mahalanobis Distance



- The Mahalanobis distance will adjust for covariance among the hormone measures at each concentration

Conceptual Example:

- Scaled and rotated Hormone A and B so that the error distribution is no longer correlated
- conc 3 is now ~4 times further away from conc 1 as conc 2

Mahalanobis distance

- To calculate the Mahalanobis distance, the response at each concentration of a test chemical was considered as a point in 11-dimensional space
  - Each axis corresponds to the natural logarithm of the measured concentration of one of the hormones included in this analysis

- Method in brief:
  1. Calculate the hormone fold-changes for each test chemical concentration compared to the DMSO control
  2. Estimate the covariance matrix that characterizes both the noise variance and correlation among hormone levels across all of the HT-H295R data
  3. Scale the computed Mahalanobis distance at each concentration of chemical screened by the number of hormones measured to give the mean Mahalanobis distance

The Mean Mahalanobis Distance (mMD)

- The mMd for a chemical at each concentration relative to the DMSO control across the steroidogenesis pathway was computed as:

$$mMd = \sqrt{(\boldsymbol{y_c} - \boldsymbol{y_1})^T \boldsymbol{\Sigma^{-1}} (\boldsymbol{y_c} - \boldsymbol{y_1})/N_h}$$

Where:

- y_c is the vector of natural log-transformed steroid hormone concentrations at the c^{th} concentration
- y_1 is the vector of natural log-transformed steroid hormone concentrations for the DMSO control
- N_h is the number of hormones with measurements for this chemical
- $\sum^{-1}$ is the estimate of the inverse covariance matrix

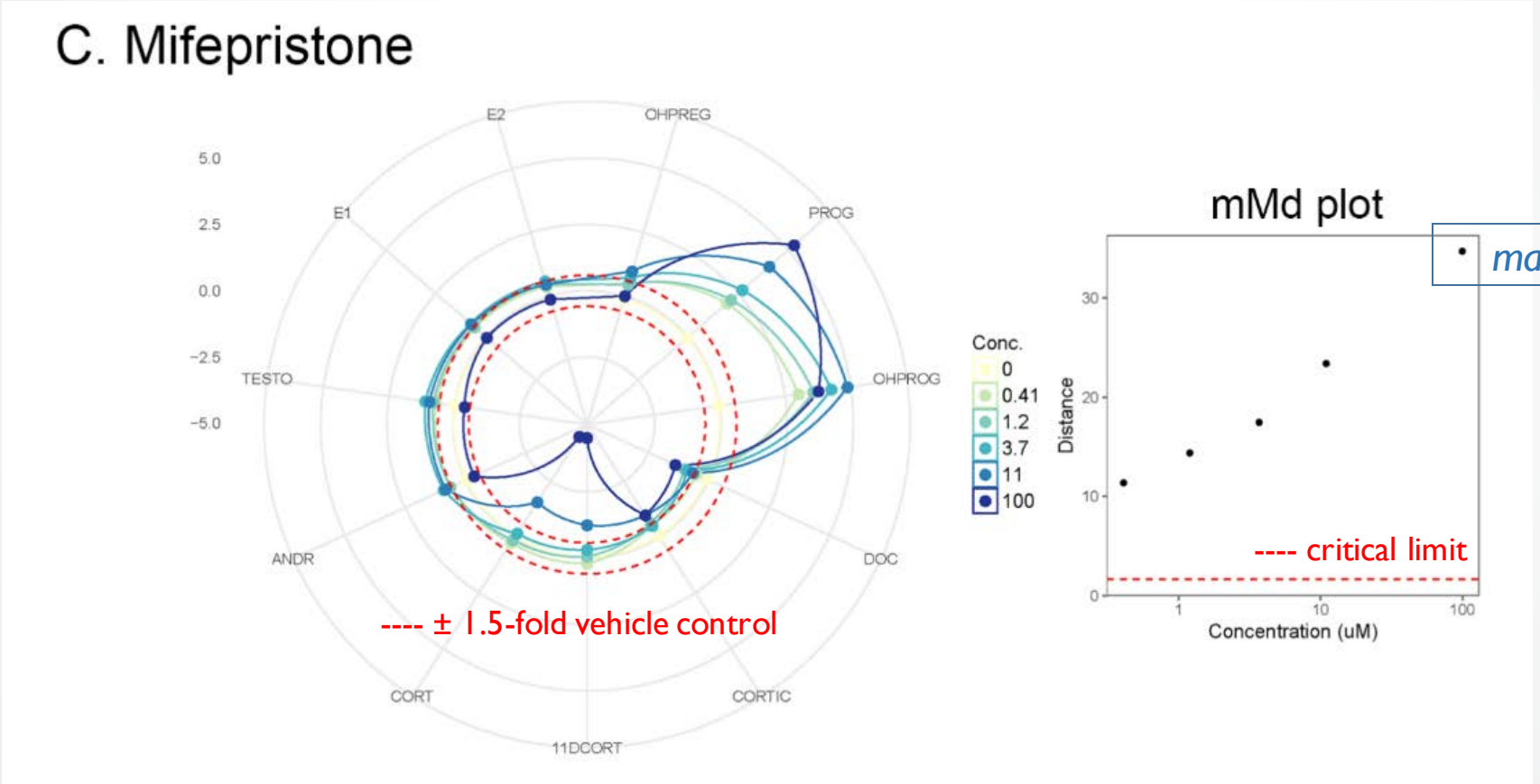
Covariance Matrix Estimation

- Fit multivariate linear model (per block) using ln-transformed hormone concentrations
- Matrix of fit residuals for data from all plates within each block were used to estimate a variance and covariance matrix
- Unweighted average of the covariance matrices across blocks = full pooled 11 X 11 covariance matrix used for the mMd calculation

Ranking Chemicals Using mMd

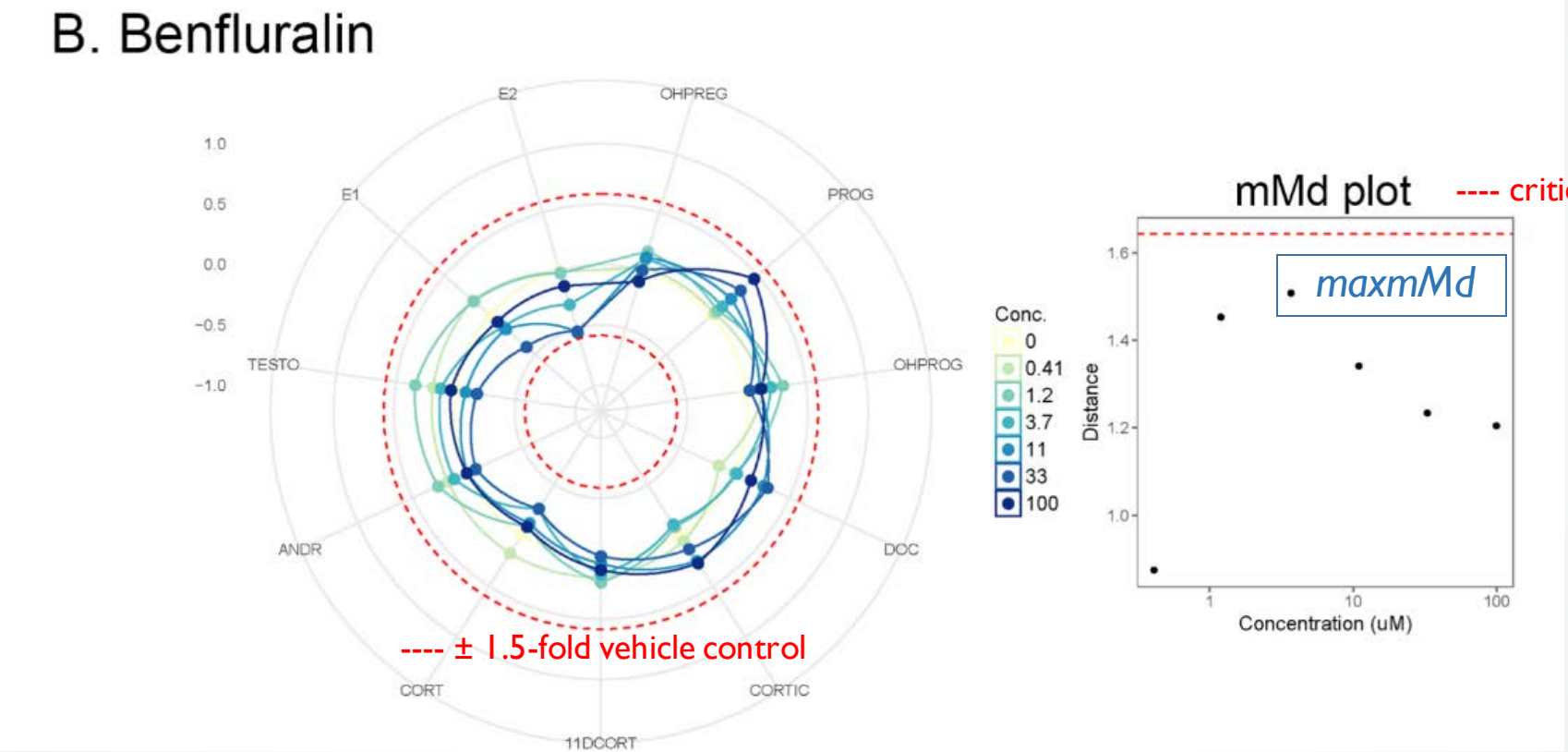
- The maximum mMd (maxmMd) is the maximum of the set of mMd values computed for all concentrations of a test chemical
  - Overall magnitude of effect of a test chemical on the steroidogenesis pathway
- As mMd generally increases with increasing concentration, a greater maxmMd should indicate:
  - Increasing concentration of chemical
  - Increased potency (i.e., activity at lower concentrations)
- Critical limit:
  - Derived to distinguish mMd values greater than what would result from noise
  - Accounts for multiple comparisons arising from comparing each concentration to the control

Example: Strong effects



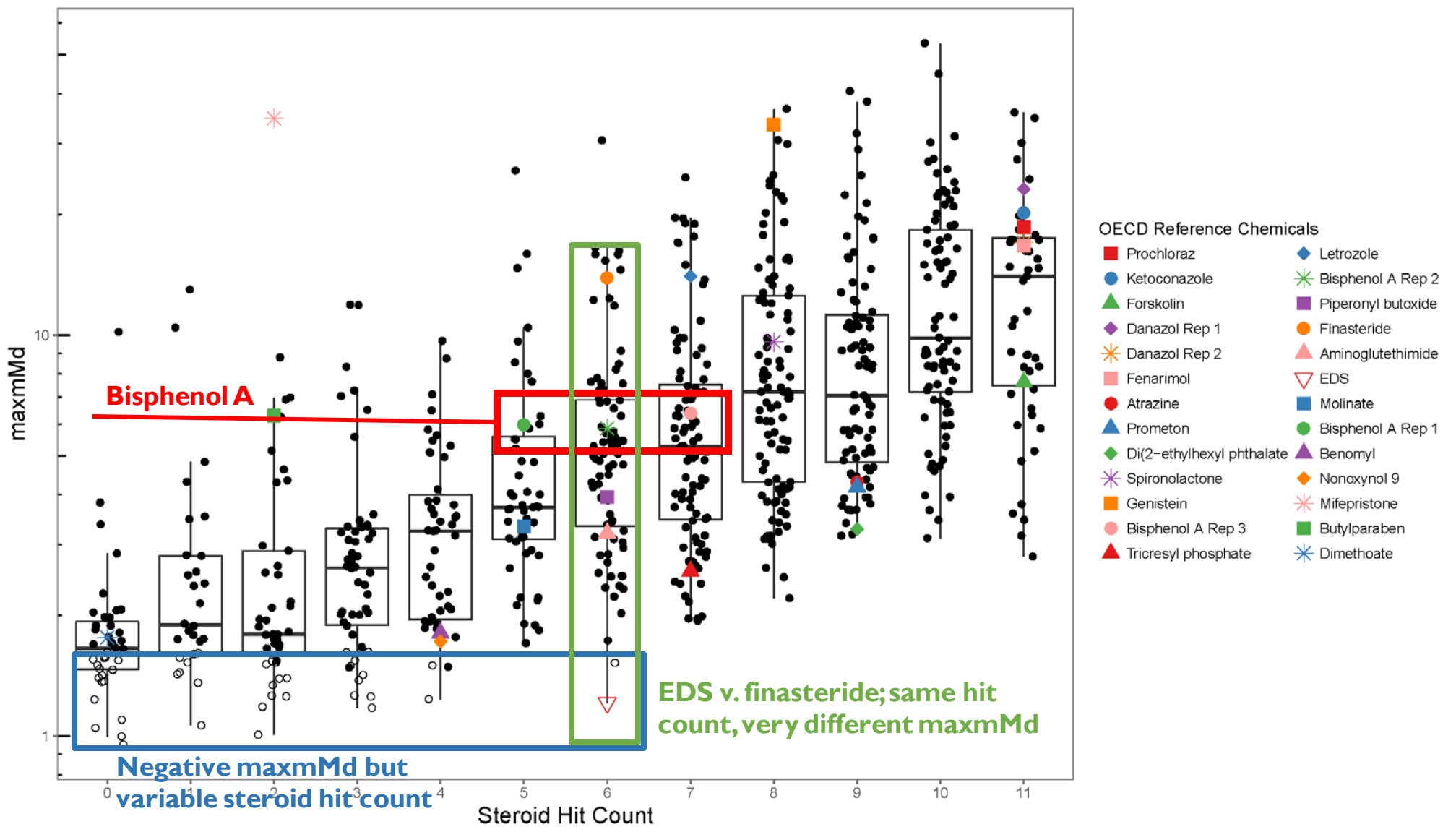
Mifepristone strongly modulated progestagens with significant effects on progesterone and OH-progesterone and moderate but non-significant trends on corticosteroids and androgens, resulting in a relatively high adjusted maxmMd of 33.

Example: Negative



Benfluralin provides an example of a chemical with a negative pathway result, with no significant concentration-response for the mMd values, as the maxmMd failed to exceed the critical limit (adjusted maxmMd of -0.14).

maxmMd was Reproducible and Quantitatively Distinguished Chemicals with Larger Effects

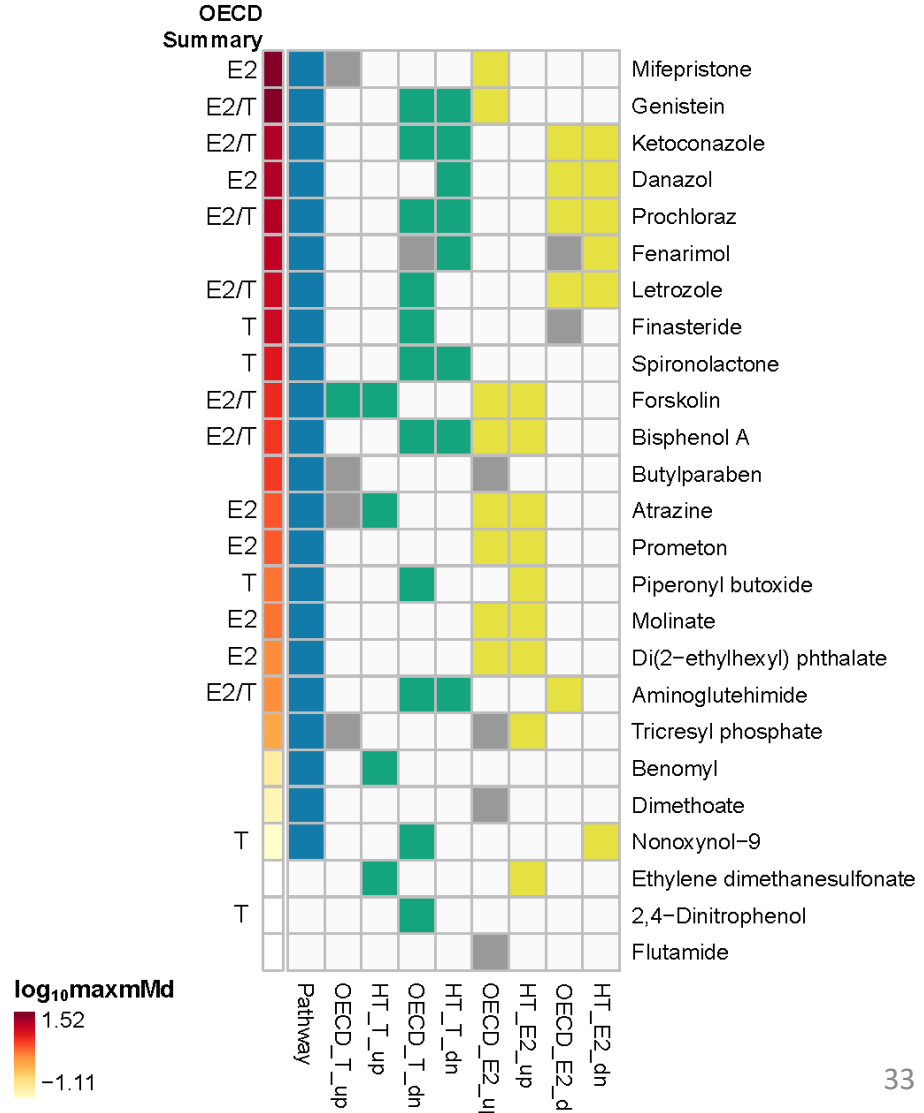


Reproducibility of the maxmMd

- 107 chemicals were replicated in > 1 block, with maxmMd ranging 1-35 for this subset
- Median maximum difference between maxmMd values across blocks ≈ 1.47 units on the arithmetic scale
- 88% of the maxmMd pathway responses replicated, with failures largely attributable to borderline activity (contrast with 65% recall for OECD ANOVA logic)

maxmMd Pathway Responses Matched the OECD Inter-laboratory Reference Chemical Activity

- Positive maxmMd pathway response (blue) was observed when signif. effects on E2 and T were observed in LT-H295R
- MaxmMd value separated strong modulators (e.g., mifepristone, prochloraz, ketoconazole, danazol, letrozole) from moderate (e.g., atrazine, molinate, di(2-ethylhexyl-phthalate) and non-active (e.g., EDS)
- Reference chemical effects on progestagen and corticosteroid biosynthesis mostly unknown



Summary and Conclusions

Evaluation of HT-H295R assay

- This detailed, performance-based comparison highlights good concordance of results, with accuracies that range 0.80 – 0.95 for effects on E2 and T
- Agreement among the labs in the inter-laboratory validation generally approached 90%
- Minor disagreement between the HT-295R and LT-H295R results occurred for chemicals with borderline activity or activity at high concentrations

maxmMd May Be Useful for Prioritization and EDSP Weight-of-evidence Applications

- Calculation of the set of mMd values reduced an 11-dimensional question to a single dimension
- Selection of the maxmMd appeared to provide a reproducible, quantitative approximation of the magnitude of effect on steroidogenesis
  - Quantitatively distinguished weak, moderate, and strong effects on one or more hormones in the pathway
- Given an mMd at each concentration, a modeled mMd at the critical limit, or the lowest concentration corresponding to a significant mMd, could be used:
  - As a concentration at which to review effects on specific hormones
  - As a lowest observable effect concentration

Acknowledgements

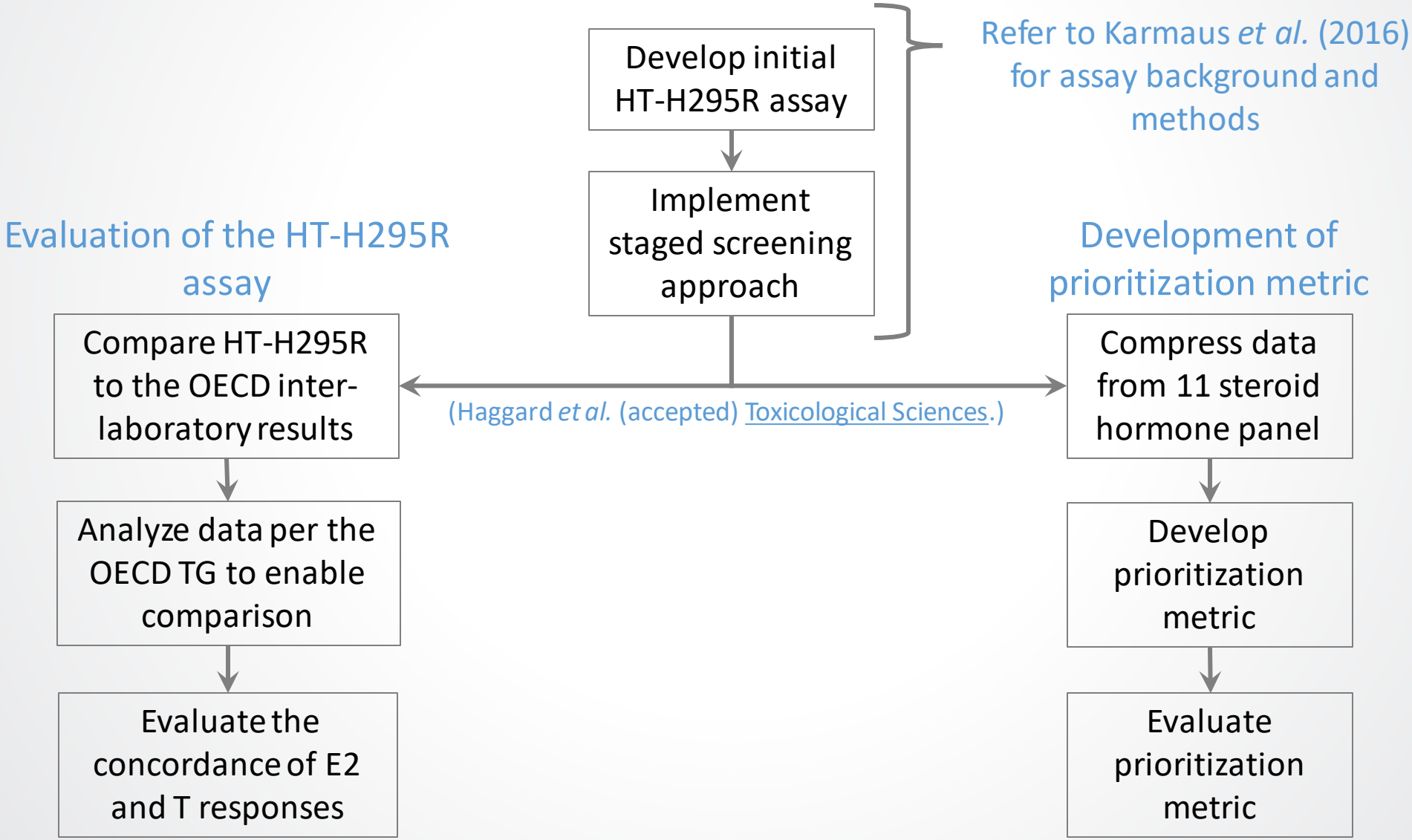
- Katie Paul Friedman (mentor)
- Woody Setzer
- Richard Judson
- Agnes Karmaus
- Matt Martin

- ORISE program
- NCCT

Questions?

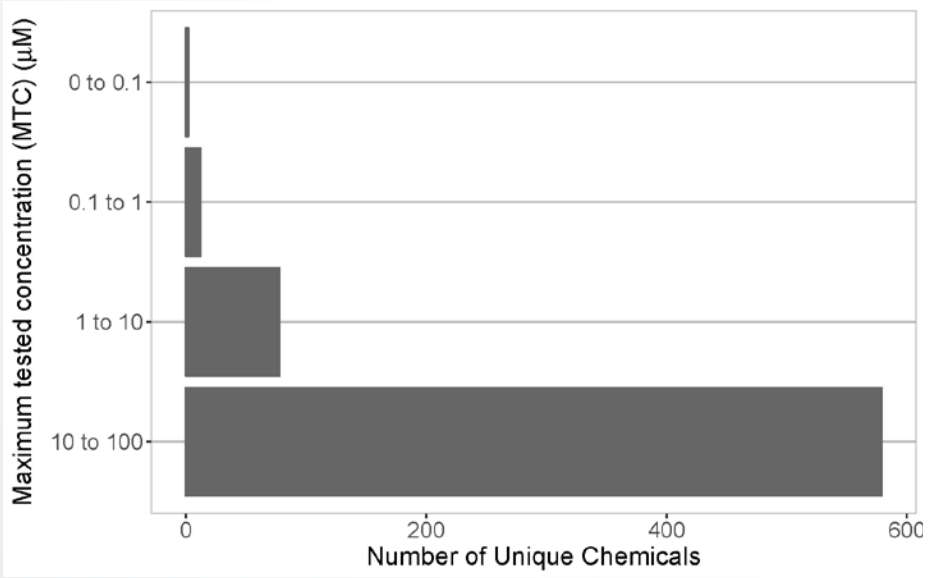
Appendix Slides

Overall Approach

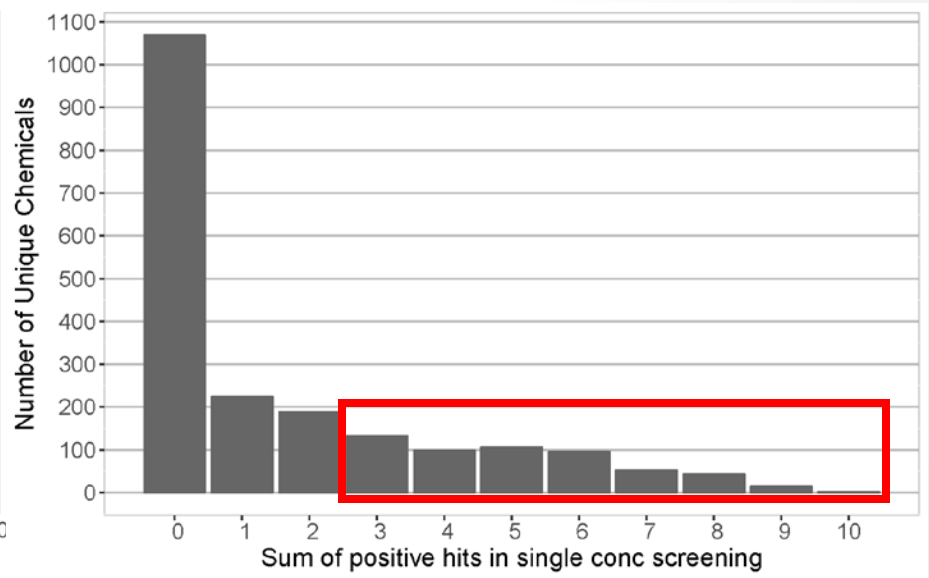


Staged Screening with HT-H295R Assay

- Maximized screening resource efficiency
- # steroid hormones affected in single concentration (along with other considerations) were used to select 656 chemicals for multi-concentration screening.



A maximum testable concentration (MTC) was determined (viability ≥ 70%).



Single concentration screening at the MTC was conducted for most of the screened space.

Karmaus *et al.* (2016) Toxicological Sciences. PMID 26781511

Brief review of covariance matrix estimation: More information

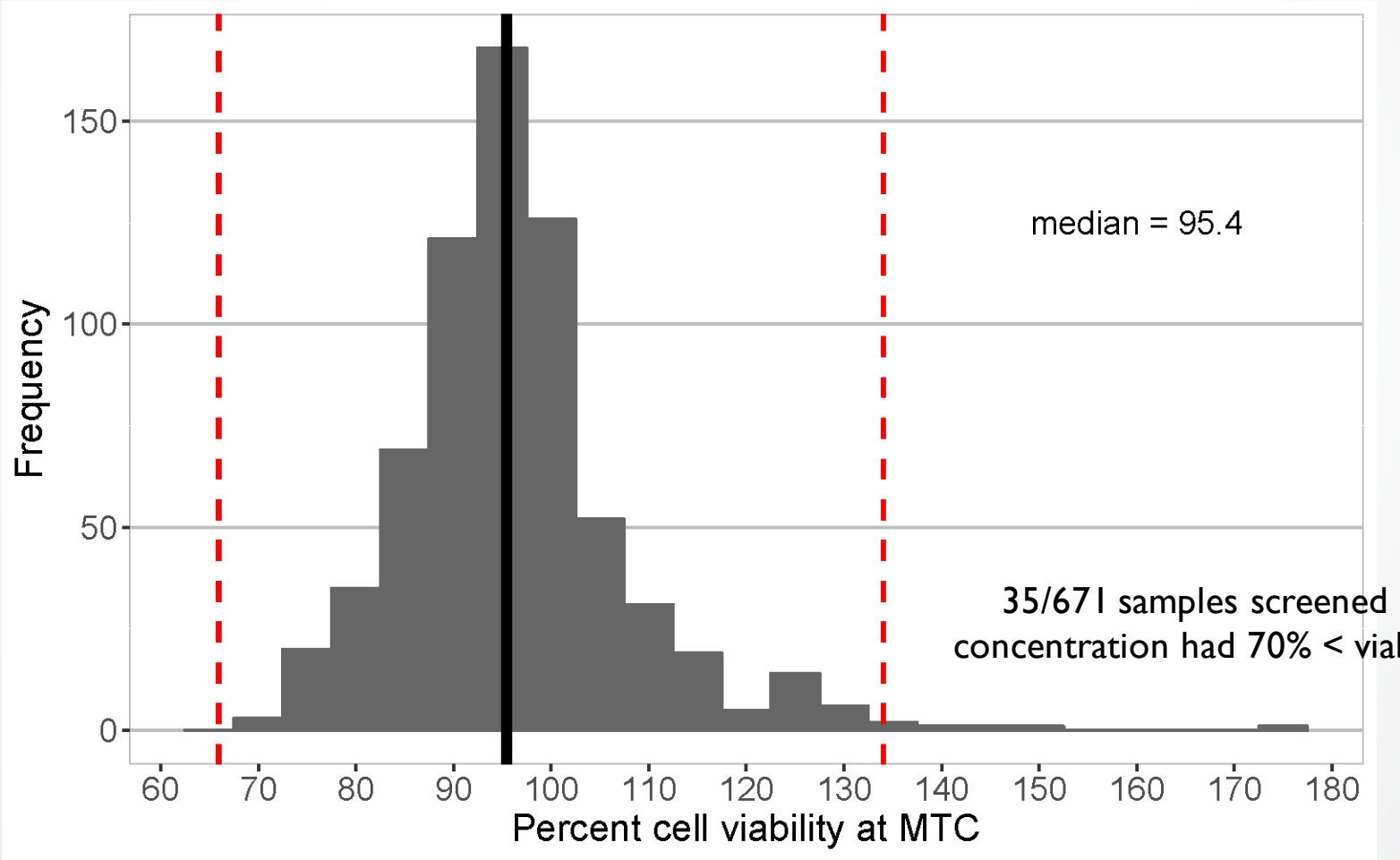
- Fit multivariate linear model (per block) using ln-transformed hormone concentrations.
- Matrix of fit residuals for data from all plates within each block were used to estimate a variance and covariance matrix.
- If any data were missing†, the hormone measure was dropped from that block prior to linear model fitting (only for 1/8 blocks, i.e. 81 chemicals, proceeded with 9/11 hormones).
- Unweighted average of the 8 block-specific covariance matrices = full pooled 11 X 11 covariance matrix used for the mMd calculation.

†The condition for missing data was "not reportable" flagged by the vendor, likely indicating a lost or "dropped" sample. During the sample analysis process, samples were flagged as "not-detected" or "not-quantifiable" when the sample was available, but the steroid hormone analyte was below the LLOQ; in such cases, a surrogate value of the LLOQ/√2 was substituted for analyses herein (CDC, 2009; Hornung and Reed, 1990). "Missing data" affected only one of the eight blocks, which contained some missing data for estrone and E2, representing 81 unique test chemicals. In this case, the computed covariance matrix for this block included only nine of the 11 steroid hormone analytes.
Removal of missing data enabled a positive definite matrix.

Critical value for positive steroidogenesis pathway results

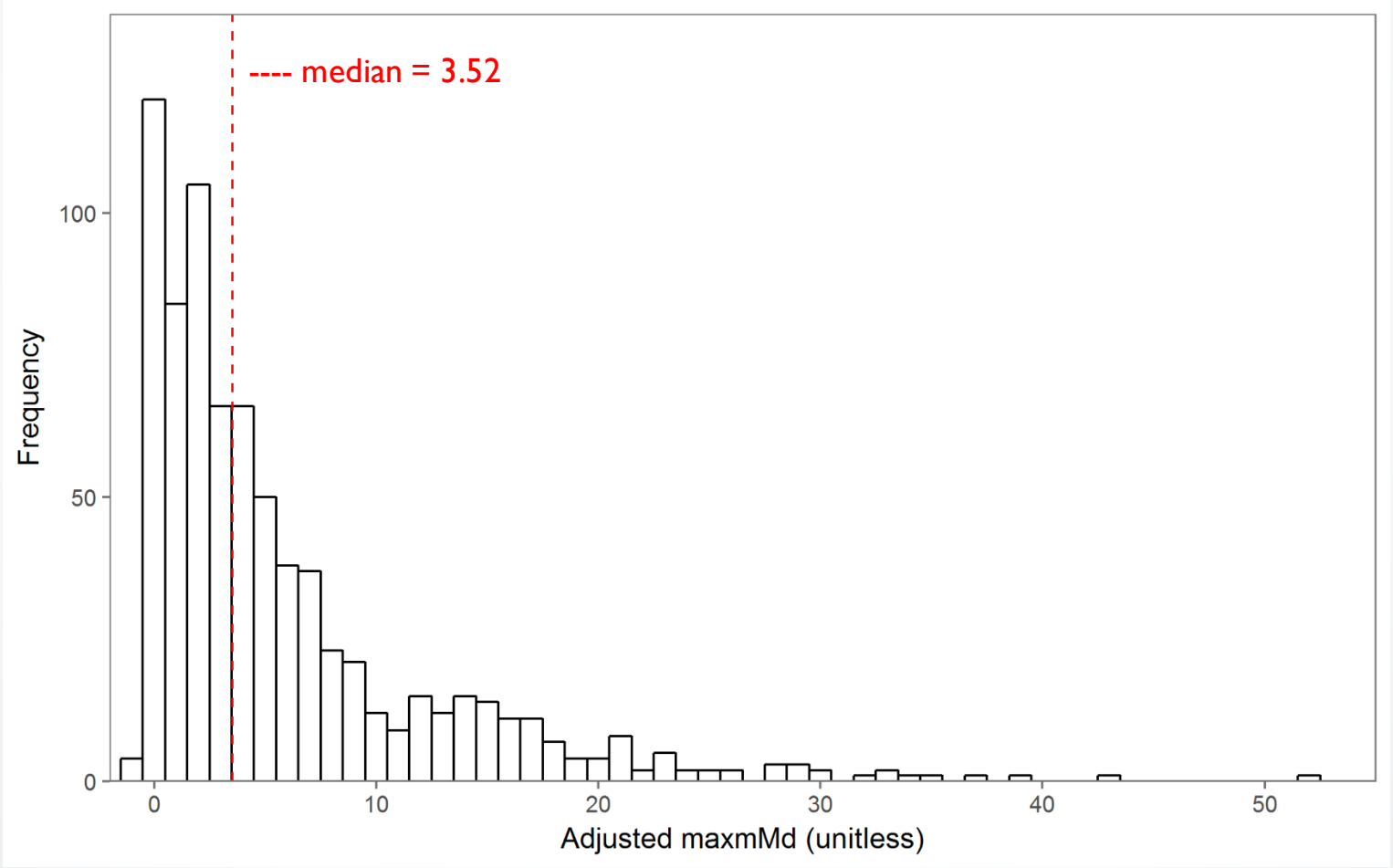
- Critical value:
  - Derived to distinguish mMd values greater than what would result from noise.
  - Accounts for multiple comparisons arising from comparing each concentration to the control.
- Similarity between mMd and Hotelling T^2
  - Hotelling T^2 used to compare two groups with multiple measures.
  - In this analysis, within-group variance-covariance matrix is used instead, using method of Nakamura and Imada (2005).
  - Analogous to adjusting for multiple comparisons for univariate tests such as the Dunnett's procedure.
- Critical value derived for approximate Type I error of 0.01 and is related to the number of hormones with data for each chemical.

Most of the MTC data corresponded to cell viability of ≥ 80%



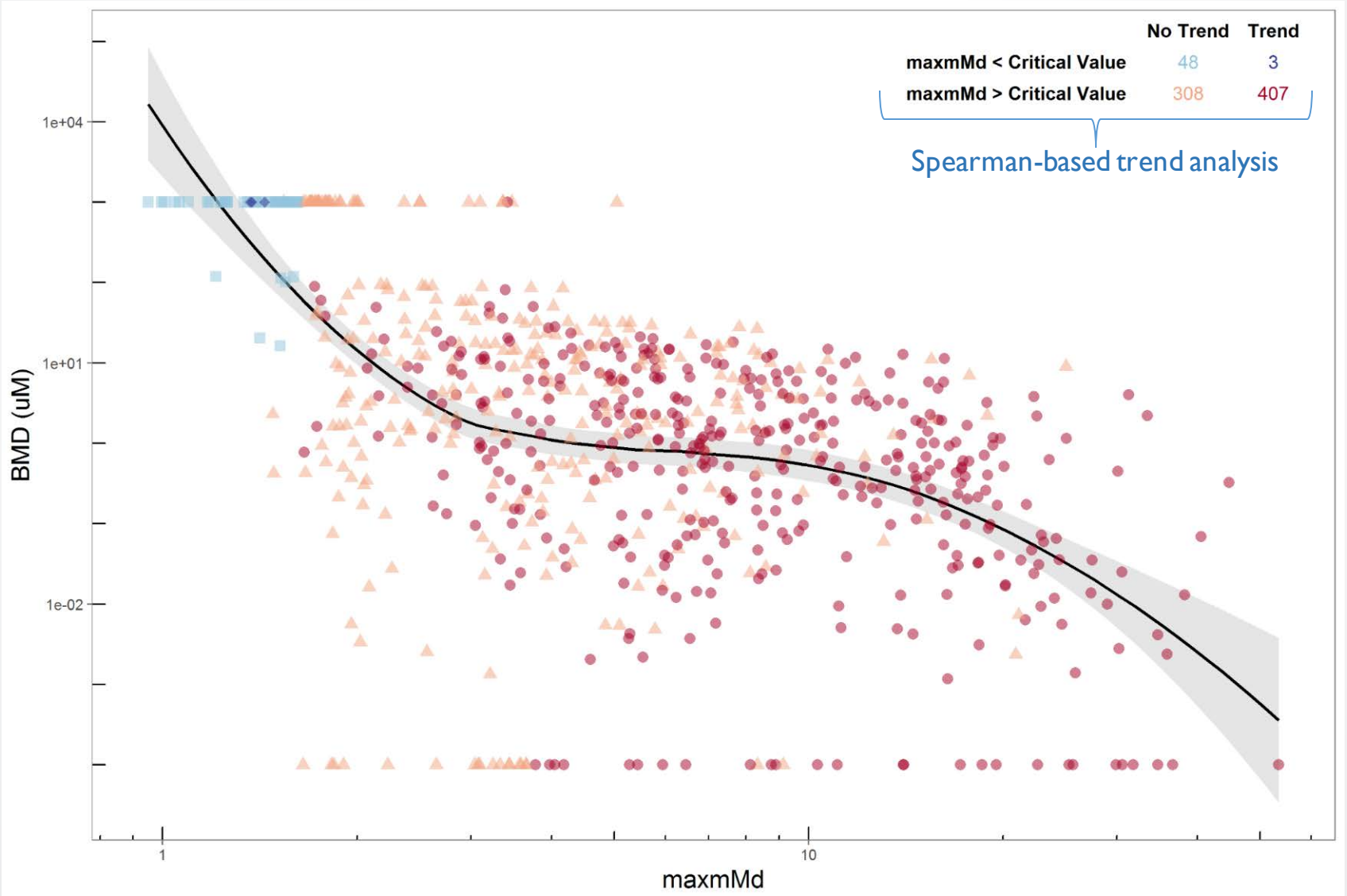
- - - ± 5* baseline median absolute deviation for the MTT assay

The maxmMd distribution for this dataset



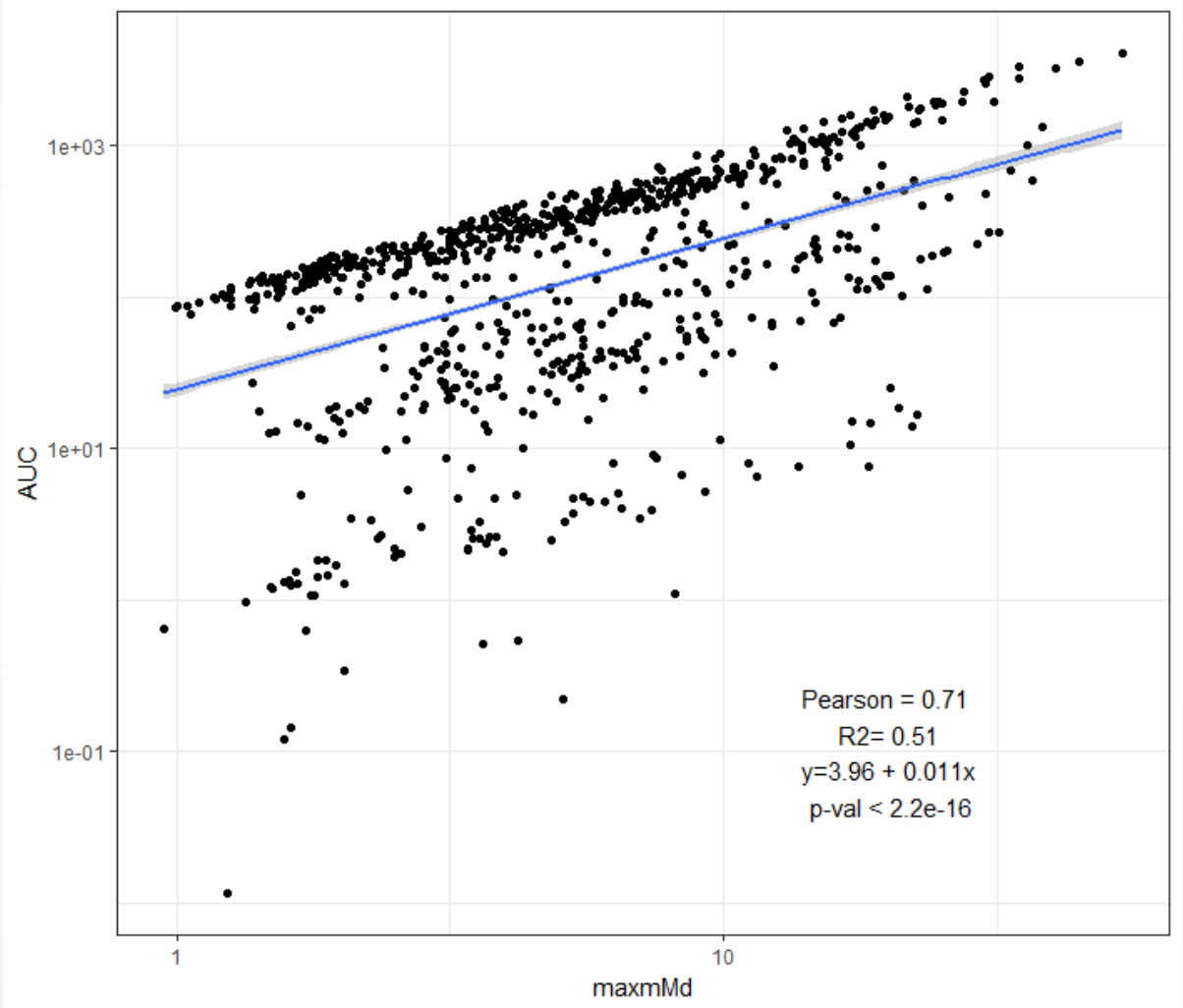
The maxmMd generally indicates potency

Concentration at which the Hill fit of the mMd data intersects with the critical value for a given chemical



The maxmMd correlates with the AUC

AUC was calculated from the mMd vs. concentration for each date-chemical-plate combination.



Limitations

- Lack of reference chemical information on the full steroidogenesis pathway
- No consideration for mitochondrial toxicity
- Potentially limited metabolic capacity of the assay
  - H295R do express xenobiotic metabolizing enzymes, but they may not generate all relevant chemical metabolites
- Current libraries are restricted to DMSO-soluble chemicals
  - Future plans include expanding chemical testing to a water-soluble library