Assessing Uncertainty in Risk Assessment Models

Eric D Watt[1,2] and Richard S Judson[2]

[1]Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee
[2]National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina

Scientific Context

- ***In vitro*** **assays are increasingly being used in risk assessments**
- **Uncertainty in assays leads to uncertainty in models used for risk assessments**

Over 2.6 million in vitro curves
- Many chemicals (> 8,000 unique)
- Many assays (> 800)

Broad assay coverage
- Numerous assay sources (> 10)
- Many biological pathways (> 400)
- Representing many species including human, rat, mouse, and fish
- Diverse detection methods including fluorescence, colorimetric, radioactive, electronic sensing, and RNA transcription

Broad chemical coverage
- Pesticides, food additives, green alternatives, endocrine reference compounds, water contaminants, fragrances, etc.

ToxCast Pipeline offers consistent analysis
- Multiple models fit to data to determine efficacy (top) and potency (AC50) (Fig 1)
- Model selection based on AIC from model fits, hit call based on efficacy relative to cutoff for winning model

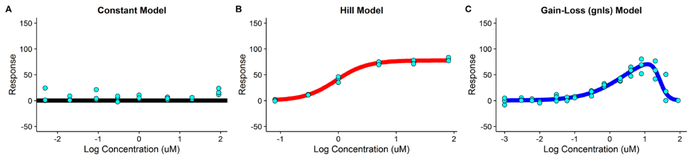


Fig 1 ToxCast models. A) Constant (cnst), B) Hill, and C) Gain-Loss (gnls) models.

Approach

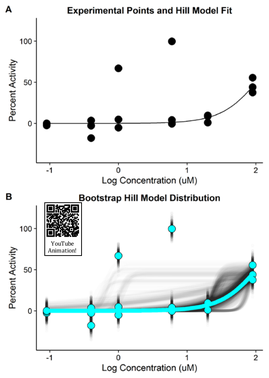


Fig 2 Bootstrap. A) Experimental response values (circles) and hill model fit. B) Uncertainty in response values and fitted model using 1000 bootstrap resamples. Experimental (cyan) and bootstrap resampled (black) response values (circles) and hill model fits (lines).

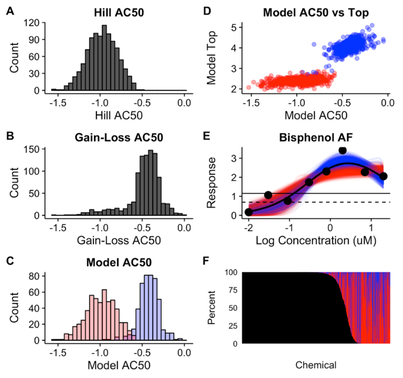


Fig 3 Bisphenol AF ATG_ERa_TRANS_up bootstrap results. Potency (AC50) values for A) Hill, B) Gnls, C) Winning model potency (hill red, gnls blue). D) Correlation between winning model efficacy (top) and AC50 (hill red, gnls blue). E) Experimental values (black circles) and model fit (black curve). Dashed line is 3x baseline median absolute deviation and solid line is assay activity cutoff. 1000 bootstrap fits are indicated (534 hill red, 466 gnls blue). F) Hit call and model selection for all chemicals. Black bars indicate chemical not a hit, red is hill model active, blue is gnls active.

Results and Applications

Case Study Name: Estrogen Receptor Model

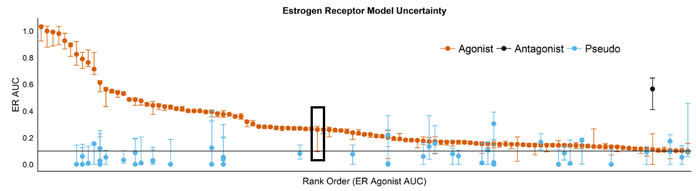


Fig 4 ER model AUC. Estrogen receptor model area under the curve (AUC) values for chemicals with an agonist AUC > 0.1. Point estimates (circles) for agonist (red), antagonist (black), and pseudo receptor (blue) where upper 95% confidence interval > 0.1.

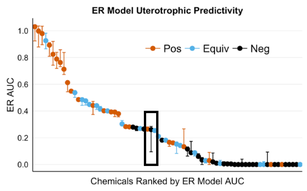


Fig 5 ER model AUC uterotrophic comparison. ER model prediction and *in vivo* guideline-like study are plotted, ranked by AUC. Color indicates *in vivo* activity of positive (red), equivocal (blue) or negative (black).

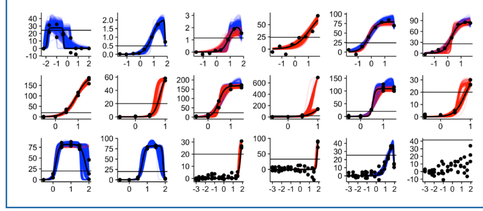


Fig 6 Exploration of source of uncertainty. Underlying assay data used to generate ER model score for Nordihydroguaiaretic (boxed Figs 4,5). Experimental points (circles), fit model (black), cutoff (horizontal line), and bootstrapped curves (hill red, gnls blue) are indicated.

- ER model recently approved for the Endocrine Disruptor Screening Program (EDSP) Tier 1 battery[1]
- Linear pathway model
- 18 ToxCast assays, 5 sources
- AUC > 0.1 predictive of activity
- Propagate bootstrap to get confidence intervals around AUC values

Bootstrap Estimates:
- **Softens activity cutoffs and improves balanced accuracy**

Case Study Name: Androgen Receptor Model

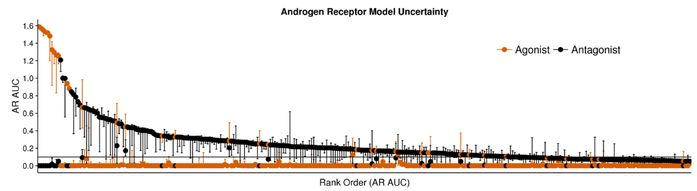


Fig 7 AR model AUC. Androgen receptor model AUC values for chemicals with an agonist or antagonist AUC > 0.05 with point estimates (circles) and 95% confidence intervals (error bars) for agonist (red) and antagonist (black) values.

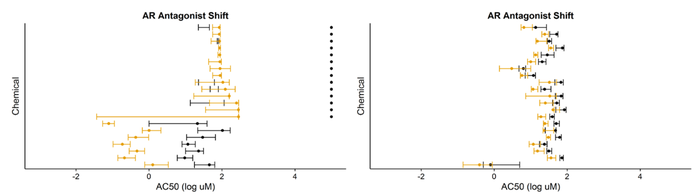


Fig 8 Androgen antagonist activity shift. Comparison of androgen antagonist assay potency with high (black) and low (gold) agonist concentration. Chemicals acting as true antagonists are expected to see a potency shift. Using bootstrap confidence intervals, we determine which chemicals have a significant potency shift and are therefore likely to be true antagonists.

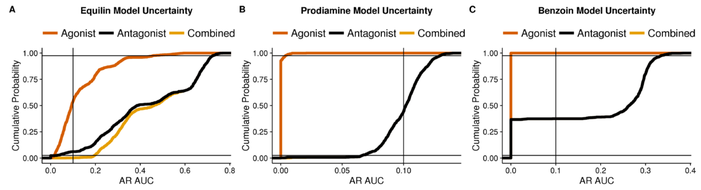
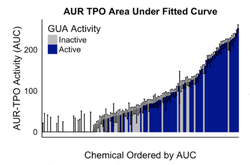


Fig 9 Androgen receptor model uncertainty distributions. Distributions of agonist (red), antagonist (black) or combined (orange) AUC values are explored using cumulative distribution function plots. A) Equilin is clearly active with some uncertainty between agonist and antagonist modes. B) Prodiamine antagonist AUC is slightly above the cutoff and narrow distribution around this value showing high confidence in the calculated score. C) Benzoin agonist and antagonist AUC point estimates are 0, but there is ~60% probability of an antagonist value in the range of 0.2-0.35, flagging this chemical as a potential false negative.

Case Study Name: Ongoing Applications



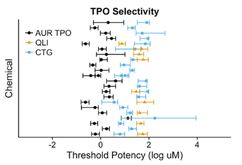
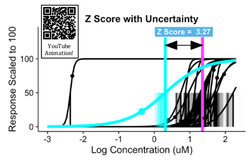


Fig 10 Thyroid peroxidase assay comparison. The AUR-TPO assay[3] is compared to the previous GUA-TPO assay. AUR activity is represented as the area under the curve of the fitted model, while historic GUA results are indicated by color. Error bars are the 95% CI in the AUR AUC calculation.

Fig 11 Thyroid peroxidase selectivity. To put the AUR-TPO hits into context, two orthogonal assays were run. QLI is a orthogonal enzyme inhibition assay and CTG tests for cytotoxicity. The 95% CIs help to confirm potency shifts between TPO and the other two assays.



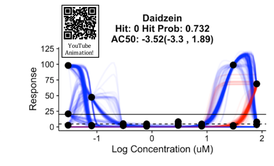


Fig 12 Cytotoxicity. The z score is calculated as the difference in potency between the assay of interest (cyan) and the median potency in cytotox assays (magenta), divided by the global median absolute deviation. Bootstrapping gives uncertainty in z score calculation.

Fig 13 Zebrafish development. Example zebrafish[2] assay chemical with wide confidence intervals. Experimental points (circles), fit model (black), cutoff (horizontal solid), 3x baseline median absolute deviation (horizontal dashed), and bootstrapped curves (hill red, gnls blue) are indicated.

Anticipated Products

- Manuscript – Methods & ER Case Study: Internal Review
- Manuscript – AR Case Study: Submitted
- Manuscript – Thyroid Peroxidase Case Study: In Prep
- Manuscript – Cytotoxicity: In Prep
- Manuscript – Zebrafish: In Prep
- R Package – Toxboot: Published
- R Package – Toxpath: In Prep
- Data to be available on dashboard and FTP

References

1. Federal Register. Use of High Throughput Assays and Computational Tools; Endocrine Disruptor Screening Program; Notice of Availability and Opportunity for Comment. 2015 Jun pp. 35350–35355. Report No.: 80 FR 35350.
2. Padilla S, Corum D, Padnos B, Hunter DL, Beam A, Houck KA, et al. Zebrafish developmental screening of the ToxCast™ Phase I chemical library. Reproductive Toxicology. 2012 Apr;33(2):174–87.
3. Paul Friedman K, Watt ED, Hornung MW, Hedge JM, Judson RS, Crofton KM, et al. Tiered High-Throughput Screening Approach to Identify Thyroperoxidase Inhibitors Within the ToxCast Phase I and II Chemical Libraries. Toxicological Sciences. 2016 May;151(1):160–80.

[1]The views expressed in this poster are those of the author[s] and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.