

Endocrine Profiling and Prioritization of Environmental Chemicals Using ToxCast™

David Reif¹, Matt Martin¹, Shirlee Tan², Keith Houck¹, Richard Judson¹, Ann Richard¹, Thomas Knudsen¹, David Dix¹, Robert Kavlock¹

¹National Center for Computational Toxicology Office of Research and Development

²Office of Science Coordination and Policy Office of Pollution Prevention, Pesticides and Toxic Substances

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY



Office of Research and Development National Center for Computational Toxicology This work was reviewed by EPA and approved for presentation but does not necessarily reflect Agency policy



Í.

Presentation outline

- I. Prioritization framework
 - **Rationale for integrated prioritization scheme**
 - ii. Definitions and notation
 - iii. Interpreting ToxScores for individual chemicals
- II. Implementation for the task of endocrine prioritization
 - i. EDSP prioritization
 - ii. Developing a prioritization scheme for EDCs
 - iii. Data sources
 - iv. Results: EDSP chemicals of interest in ToxCast Phase-I
 - v. Results: empirical distribution of EDSP chemicals of interest
 - vi. Results: guidepost ("spike-in") chemicals
 - vii. Results: exploring ToxScores in the context of in vivo results
 - viii. Results: rank by specific slices (e.g. AR) in the context of in vivo results
 - ix. Results: chemical classes
 - x. Results: simulation studies assess sensitivity to spurious assay results
- III. Future directions
 - i. Alternative implementations
 - ii. Incorporation of new/other data (e.g. QSAR models and other extant tools)
- **IV.** Conclusions



Rationale for an integrated chemical prioritization scheme





A numerical index that can be used for ranking (instead of absolute thresholds) is more flexible for different prioritization tasks and can better accommodate new data, new chemicals, data adjustments, etc.



Definitions & Notation

$$\mathsf{ToxScore} = \sum_{1}^{I} \mathsf{w}_{i} * \mathsf{assay}_{i} + \sum_{1}^{C} \mathsf{w}_{c} * \mathsf{chemProp}_{c} + \sum_{1}^{P} \mathsf{w}_{p} * \mathsf{pathway}_{p}$$

Component: Individual in-vitro assays, chemical properties/descriptors, etc. (e.g. ERbinding assay from Novascreen)

Slice: "Pie" slices representing individual *components* or aggregrations of multiple related *components* (e.g. slice *i*=2 represents multiple in-vitro assays related to the estrogen receptor)

Domain/Axis: Domain/field of knowledge; represented by the slice(s) of a given color family (e.g. all chemical properties have slices in some shade of orange)

Example Sentence:

"For each chemical, the ToxScore[™] integrates information across multiple *domains*, which are composed of one or more *slices*, which are composed of one or more *components*."





Interpreting ToxScores for individual chemicals

profile/ Each chemical signature/ gives a score index (ToxScore) used for ranking chemicals fingerprint

ToxScore = f(In vitro assays + Chemical properties + Pathways)

$$\mathsf{ToxScore} = \sum_{1}^{I} \mathsf{w}_{i} * \mathsf{assay}_{i} + \sum_{1}^{C} \mathsf{w}_{c} * \mathsf{chemProp}_{c} + \sum_{1}^{P} \mathsf{w}_{p} * \mathsf{pathway}_{p}$$





Interpreting ToxScores for individual chemicals





EDSP prioritization



EPA published the <u>draft list of initial pesticide active ingredients and pesticide inerts</u> to be considered for screening under the Federal Food, Drug and Cosmetic Act for public notice and comment in a <u>2007 Federal Register Notice</u> [PDF file, 18pp., 131KB, <u>About PDF</u>]. The draft list was produced using the approach described in the September 2005 Federal Register Notice [PDF file, 17pp., 125KB, <u>About PDF</u>], and includes chemicals that the Agency has decided should be tested first, based upon exposure potential. <u>How to comment</u>.



EDSP prioritization



Priority Setting - Approach for Initial Screening

The approach used by EPA for selecting 50 to 100 chemicals for initial screening under the Federal Food, Drug and Cosmetic Act is summarized below. Nothing in the approach for selecting the initial list would provide a basis to infer that any of the chemicals selected for the list interferes with or is suspected to interfere with the endocrine systems of humans or other species. This action may be of interest to those who are involved with, or are interested in, pesticide chemicals or the topic of endocrine disruptors.

The approach includes consideration of the most current databases and priority-setting tools available. For this approach EPA:

- 1. Focused chemical selection for this initial list on the subset of chemicals for which testing is required (i.e., pesticide chemicals);
- 2. Used exposure data as the primary basis for chemical selection;
- 3. Deferred consideration of nominations from the public;
- 4. Excluded mixtures; and
- 5. Excluded chemicals that are no longer produced or used in the United States.

The approach described in the September 2005 Federal Register notice further indicated that the following would be excluded from the list of chemicals for initial screening.

- 1. Substances anticipated to have low potential to cause endocrine disruption (e.g., most polymers with number average molecular weight greater than 1,000 daltons, strong mineral acids, and strong mineral bases);
- 2. "Positive control" chemicals used by EPA for the validation of the screening assays proposed for the Tier 1 battery.

Initial List of Chemicals

EPA published the <u>draft list of initial pesticide active ingredients and pesticide inerts</u> to be considered for screening under the Federal Food, Drug and Cosmetic Act for public notice and comment in a <u>2007 Federal Register Notice</u> [PDF file, 18pp., 131KB, <u>About PDF</u>]. The draft list was produced using the approach described in the September 2005 Federal Register Notice [PDF file, 17pp., 125KB, <u>About PDF</u>], and includes chemicals that the Agency has decided should be tested first, based upon exposure potential. <u>How to comment</u>.



EDSP prioritization



screening to eval

uate w	/hether the program could be improved or optimized.	
	1. Substances anticipated to have low potential to cause endocrine disruption (e.g., most porymers with humber average molecular weight greater than 1,000 dattors, strong minimeral actos, and strong mineral bases);	• ••
	2. "Positive control" chemicals used by EPA for the validation of the screening assays proposed for the Tier 1 battery.	
	EPA proposed the approach in a previously published December 30, 2002 Federal Register Notice [PDF file, 19pp., 119Ke, Loeut PDF] and received comments on it.	
	Subsequent Approaches for Chemical Selection	
	EPA anticipates that it may modify its chemical selection approach for subsequent screening lists based on experience gained from the results of testing of chemicals on the initial list, the need for a broader approach in the future to incorporate different categories of chemicals (e.g., non-pesticide substances) and additional pathways of exposure, and the availability of new priority-setting tools (e.g., <u>High Throughput Pre-Screening (HTPS) or Quantitative Structure Activity Relationship (QSAR) models</u>). In addition, the Agency intends to conduct a review of the data received from the screening to evaluate whether the program could be improved or optimized.	
	Initial List of Chemicals	

EPA published the draft list of initial pesticide active ingredients and pesticide inerts to be considered for screening under the Federal Food, Drug and Cosmetic Act for public notice and comment in a 2007 Federal Register Notice [PDF file, 18pp., 131KB, About PDF]. The draft list was produced using the approach described in the September 2005 Federal Register Notice [PDF file, 17pp., 125KB, About PDF1, and includes chemicals that the Agency has decided should be tested first, based upon exposure potential. How to comment,



Developing a prioritization scheme for EDCs

ToxScore = f(In vitro assays + Chemical properties + Pathways)



The ToxScore index is calculated from a weighted combination of all data sources for each chemical.

For each slice, distance from the origin (center) is proportional to the normalized value (e.g. assay potency or predicted permeability) of the component data points comprising that slice, and the width (in radians) indicates the relative weight of that slice in the overall ToxScore calculation.

The slices are drawn counter-clockwise from the right, so in this example, the AR slice is #1, the ER slice is #2, etc.



CPCP Meeting

17 December 2009

Data sources: ToxCast in vitro HTS assays



Cellular Assays

- Cell lines
 - HepG2 human hepatoblastoma
 - A549 human lung carcinoma
 - HEK 293 human embryonic kidney

Primary cells

- Human endothelial cells
- Human monocytes
- Human keratinocytes
- Human fibroblasts
- Human proximal tubule kidney cells
- Human small airway epithelial cells

Biotransformation competent cells

- Primary rat hepatocytes
- Primary human hepatocytes

Assay formats

- Cytotoxicity
- Reporter gene
- Gene expression
- Biomarker production
- High-content imaging for cellular phenotype

Biochemical Assays

- Protein families
 - GPCR
 - NR
 - Kinase
 - Phosphatase
 - Protease
 - Other enzyme
 - Ion channel
 - Transporter
- Assay formats
 - Radioligand binding
 - Enzyme activity
 - Co-activator recruitment

[Slide adapted from Keith Houck]



IN VITRO ASSAYS

Data sources



> AR

"ATG AR TRANS" "NCGC AR Agonist" "NVS NR rAR" "NVS NR hAR" "NCGC AR Antagonist"

> ER

"NVS_NR_hER" "NVS_NR_bER" "NCGC_ERalpha_Agonist" "ATG_ERa_TRANS" "ATG_ERE_CIS" "NCGC_ERalpha_Antagonist"

> TR

"NCGC TRbeta Agonist" "ATG THRaI TRANS" "NVS NR hTRa" "CLZD UGTIAI 48" "NCGC TRbeta Antagonist"

> Other XME/ADME

"CLZD CYPIAI 48" "CLZD CYPIA2 48" "CLZD CYP2B6 48" "CLZD CYP3A4 48" "NVS ADME hCYPIAI" "NVS ADME hCYPIA2" "NVS ADME hCYPIBI" "NVS ADME hCYP2A6" "NVS ADME hCYP2B6" "NVS ADME hCYP2C18" "NVS ADME hCYP2C19" "NVS ADME hCYP2C19 Activator" "NVS ADME hCYP2C8" "NVS ADME hCYP2C9" "NVS ADME hCYP2D6" "NVS ADME hCYP2EI" "NVS ADME hCYP2]2" "NVS ADME hCYP3A4" "NVS ADME_hCYP3A5" "NVS ADME hCYP4F12 Activator" "NVS ADME rCYP1A1" "NVS ADME hCYP4FI2" "NVS ADME rCYPIA2" "NVS ADME rCYP2A1" "NVS ADME rCYP2A2" "NVS ADME rCYP2B1" "NVS ADME rCYP2CII" "NVS ADME rCYP2C12" "NVS ADME rCYP2C13" "NVS ADME rCYP2C6" "NVS ADME rCYP2D1" "NVS ADME rCYP2D2" "NVS ADME rCYP2EI" "NVS ADME rCYP3A1" "NVS ADME rCYP3A2" "CLZD SULT2AI 48" "CLZD HMGCS2 48" "NVS ADME hCYPI9AI"

> Other NR

"ATG Ahr CIS" "ATG CAR TRANS" "ATG ERRa TRANS" "ATG ERRg TRANS" "ATG FXR TRANS" "ATG GR TRANS" "ATG GRE CIS" "ATG LXRa TRANS" "ATG LXRb TRANS" "ATG PPARa TRANS" "ATG PPARd TRANS" "ATG PPARg TRANS" "ATG PXR TRANS" "ATG PXRE CIS" "ATG RARa TRANS" "ATG RARb TRANS" "ATG_RARg_TRANS" "ATG_RXRa_TRANS" "ATG_RXRb_TRANS" "NCGC LXR Agonist" "NCGC_PPARg_Agonist" "NCGC PXR Agonist human" "NCGC PXR Agonist rat" "NCGC RXRa Agonist" "NVS NR bPR" "NVS NR hCAR Agonist" "NVS NR hFXR" "NVS NR hCAR" "NVS NR hGR" "NVS NR hPPARa" "NVS NR hPPARg" "NVS NR hPR" "NVS NR hPXR" "NVS NR hRAR"

For a complete description of all data sources and links to data, see: Judson et al. (2009) *Environ Health Perspect*



Data sources



CHEMICAL PROPERTIES

> LogP_TPSA "LogP_TPSA" > Predicted CaCO-2 "PCaco_QP"

PATHWAYS

> KEGG pathways

"PS_KEGG_Adipocytokine_signaling_pathway"	"PS_KEGG_Androgen_and_estrogen_metabolism"
"PS_KEGG_Biosynthesis_of_steroids"	$"PS_KEGG_Biosynthesis_of_steroids_Mus_musculus"$
"PS_KEGG_GnRH_signaling_pathway_Rattus_norveg	gicus" "PS_KEGG_Insulin_signaling_pathway"
"PS_KEGG_Melanogenesis_Rattus_norvegicus"	"PS_KEGG_PPAR_signaling_pathway"

> Ingenuity pathways

 "PS_Ingenuity_Aryl_Hydrocarbon_Receptor_Signaling"
 "PS_Ingenuity_Estrogen_Receptor_Signaling"

 "PS_Ingenuity_Insulin_Receptor_Signaling"
 "PS_Ingenuity_PPARaRXRa_Activation"

 "PS_Ingenuity_RAR_Activation"
 "PS_Ingenuity_TRRXR_Activation"

> Disease classes

"PS_Disease_Goh_Endocrine" "PS_Disease_Goh_Developmental"

"PS_KEGG_Androgen_and_estrogen_metabolism_Mus_musculus" "PS_KEGG_GnRH_signaling_pathway" "PS_KEGG_Melanogenesis" "PS_KEGG_Thyroid_cancer"

"
"PS_Ingenuity_Glucocorticoid_Receptor_Signaling"
"PS_Ingenuity_PPAR_Signaling"

For a complete description of all data sources and links to data, see: Judson et al. (2009) *Environ Health Perspect*



EDSP chemicals of interest in ToxCast Phase-I



The chemicals are arranged in alphabetical order (by name)

EDSP chemicals show a range of activities across the ToxScore components



Of endocrine interest, but not in the official Tier-I screening list





EDSP chemicals of interest in ToxCast Phase-I



17 December 2009



Distribution of EDSP chemicals of interest

(sorted by ToxScore)

path





Distribution of EDSP chemicals of interest

(sorted by ToxScore)



The ToxScore (horizontal axis) for each chemical (vertical axis) is symbolized by a box, sorted according to overall ToxScore.

EDSP chemicals of interest are highlighted (solid red boxes) along the sorted ToxScore distribution for all 309 ToxCast Phase-I chemicals.





Guidepost ("spike-in") chemicals





Data sources: Multigeneration Reproductive (MGR) studies captured in ToxRefDB



CPCP Meetin 17 December

[Slide adapted from Matt Martin]



Exploring ToxScores in the context of in-vivo results

ToxScore = f(In vitro assays + Chemical properties + Pathways)

ToxScores without an in-vivo domain can be annotated according to ToxRefDB endpoint(s), but the in vivo domain does <u>not</u> contribute to the ToxScore calculation



Here, the in vivo aggregated endpoints include multigenerational study effects in endocrine organs, the reproductive tract, offspring survival, reproductive outcome and performance, and sexual developmental landmarks (e.g. PPS, VO, AGD)



Rank by specific slices (top 10% AR, ER, or TR) in the context of in vivo results





While the in vivo endpoints from multi-gen studies do not necessarily reflect endocrine disruption, potent in vitro assay hits warrant further inspection of the chemicals involved.

Different in vitro assays have differing levels of association with in vivo endpoints.

The in vitro assays may provide new information on biological mechanisms that are not targeted by current in vivo studies.

There are many chemicals for which there are no in vivo test results for particular endpoints



Chemical classes: ToxScore plots for all **Triazoles**



For any (sub)class of chemicals, the profiles are informative for comparing and contrasting class members.

For example, the Triazoles show:

Similar LogP_TPSA scores Similar Ingenuity pathway scores Similar Other XME/ADME scores

Different AR, ER, and TR scores



Cyproconazole

Fenbuconazole



Flusilazole



Difenoconazole



Diniconazole

Hexaconazole

Propiconazole



Myclobutanil





Triticonazole

Tetraconazole

Triadimefon

Triadimenol



Assessing the stability of ToxScore rankings in the presence of spurious assay results



Methods: The simulated probability of a spurious result on a given component assay varies from 5% - 20% (NOTE: results for *chemProp* slices were held constant because they do not represent stochastic assays). Each colored data point in the figure shows the mean simulated ToxScore under each condition.

Results: While the absolute *value* of ToxScores may change, the relative ranks of chemicals are generally preserved. In situations where a chemical's absolute rank changes, it tends to swap positions with a neighbor. This is in contrast to the large shifts in rank that would occur in a prioritization scheme reliant on singular pieces of information, wherein individual errors would shift chemicals between entire priority regions (e.g. a chemical assigned a top quartile priority rank is shifted to the bottom quartile).



Í.

Presentation outline

- I. Prioritization framework
 - **Rationale for integrated prioritization scheme**
 - ii. Definitions and notation
 - iii. Interpreting ToxScores for individual chemicals
- II. Implementation for the task of endocrine prioritization
 - i. EDSP prioritization
 - ii. Developing a prioritization scheme for EDCs
 - iii. Data sources
 - iv. Results: EDSP chemicals of interest in ToxCast Phase-I
 - v. Results: empirical distribution of EDSP chemicals of interest
 - vi. Results: guidepost ("spike-in") chemicals
 - vii. Results: exploring ToxScores in the context of in vivo results
 - viii. Results: rank by specific slices (e.g. AR) in the context of in vivo results
 - ix. Results: chemical classes
 - x. Results: simulation studies assess sensitivity to spurious assay results
- III. Future directions
 - i. Alternative implementations
 - ii. Incorporation of new/other data (e.g. QSAR models and other extant tools)
- **IV.** Conclusions



Alternative implementations







Prioritization tasks might:

A) Incorporate additional components (slices) that may be from other domains (e.g. Consideration of exposure potential);

B) Customize individual domains (e.g. Add a targeted set of chemical descriptors);

C) Adjust weighting schemes according to specific prioritization tasks or component (slice) meaning (e.g. The weights $(w_{i=1,2,3})$ of In vitro assay slices 1, 2, and 3 (representing AR, ER, and TR, respectively) have been increased).



Incorporation of new/other data



The framework is amenable to incorporation of new data from the EDSP Tier-I battery, subsequent phases of ToxCast, and other data from any number of sources.

For example:

QSAR models, such as the Mid-Continent Ecology Division's system for ER binding potential, could be incorporated into the *chemProp* domain.

Toxicological knowledgbases, such as the Endocrine Disruptor Priority Setting Database, could be used to assign priors to particular chemicals.

Exposure tools, such as ExpoCast, could provide data for the *exposure* domain.

Predictive signatures, such as those developed as part of ToxCast, could be added as components.

... and many more



Conclusions

This work was reviewed by EPA and approved for presentation but does not necessarily reflect Agency policy

This implementation indicates that an integrated approach, wherein multiple domains of toxicological knowledge are simultaneously incorporated into chemical prioritization, gives a reasonably stable priority rank across the ToxCast Phase-I chemicals.

The inclusion of benchmark chemicals (akin to a "spike-in" set) as internal controls reduces the probability that potentially hazardous chemicals will be improperly assigned low priority for further testing and makes this a promising approach for diverse chemical prioritization tasks.

The ToxRefDB in vivo results may be useful for evaluation of other, specific prioritization tasks.

The framework developed here provides graphical insight into the multiple domains considered in chemical profiling and prioritization.

It is amenable to incorporating extant prioritization schemes and relevant data from diverse sources, thereby facilitating meta-analysis across Agency resources.

Because ToxScores are intended for relative ranking, particular implementations of this framework can be continually updated with new chemicals and future data.