Report on Peer Review of Draft NTP Approach to Genomic Dose-Response Modeling

Scott S. Auerbach, Ph.D.
Biomolecular Screening Branch
National Institute of Environmental Health Science

Board of Scientific Counselors Meeting
December 8, 2017
Genomic Dose Response

What is it?

In vitro
In vivo
X multiple dose levels

Measure Gene Expression (Affy, S1500+, RNA-seq, etc)

Identify genes responding to treatment

Fit dose response models to genes and identify potency of response

Map genes to curated gene sets

Identify gene sets that are “active”

Determine potency of “active” gene sets

Report the potency of the most sensitive gene set as the biological effect point of departure (BEPOD)

Why do we care?

Can quickly query a wide swath of biological space to identify an effect level that approximate the potency of traditional toxicological endpoints
NTP is pursuing in vivo and in vitro transcriptomics in dose-response format for determining screening level biological potency and for identification of molecular processes that are altered by test articles.

Data from these studies are intended to support margin of exposure based assessments that can help in prioritization and setting interim exposure limits.

Expert panel meeting was held to review the NTP’s proposed approach to generating and analyzing the data from genomic dose-response studies (GDRS).
The Meeting
National Toxicology Program Proposed Approach to Genomic Dose-Response Modeling

Introduction

Changes in the approach to toxicological assessment [1] and the advent of inexpensive, high-throughput transcriptomics data generation platforms have led to significant interest in the integration of genomic dose-response studies into the hazard characterization and risk assessment process [2]. Currently, many questions exist regarding how best to design, perform, and interpret genomic dose-response studies in a manner that most effectively facilitates integration of these types of assessments into the hazard characterization paradigm. Consensus or, at a minimum, guidance on how to carry out and analyze these types of studies would be helpful in advancing the use of their findings in toxicology and risk assessment decision making.

This document describes a proposed framework for performing genomic dose-response analysis that is largely consistent with published approaches. The described approach is targeted at developing screening-level hazard assessments of test articles that can be used for prioritization and the setting of interim exposure limits, particularly for in vivo studies.

NTP is convening an expert panel on October 23–25, 2017, at the National Institute of Environmental Health Sciences, Research Triangle Park, NC to obtain input on specific details of its proposed approach to genomic dose-response modeling. NTP will carefully consider the panel’s input and recommendations in determining what changes to the approach might be needed prior to finalization. NTP will also continue to monitor the scientific literature with regard to the development of improved approaches to data modeling and analysis. Importantly, in reviewing the proposed approach, we ask readers to keep in mind that NTP’s goal in pursuing genomic dose-response studies is to quickly and cost effectively develop sensitive, screening-level potency estimates for test articles and provide a degree of contextualization to facilitate qualitative interpretation of observed genomic changes.

A Method to Integrate Benchmark Dose Estimates with Genomic Data to Assess the Functional Effects of Chemical Exposure

Russell S. Thomas,¹,¹ Bruce C. Allen,† Andy Nong,⁎ Longlong Yang,⁎ Edilberto Bermudez,⁎ Harvey J. Clewell III,⁎ and Melvin E. Andersen⁎

Software
BMDExpress: a software tool for the benchmark dose analyses of genomic data
Longlong Yang¹, Bruce C Allen² and Russell S Thomas⁎¹

Fred Parham
Mary Wolfe
• Webinars
  – **Goal:** Bring everyone up to speed before the meeting
  – Scott Auerbach (NTP)
    • The NTP Proposed Approach to Genomic Dose-Response Modeling (August 30, 2017)
  – Lyle Burgoon (US Army)
    • Overview of the U.S. Army Approach to Genomic Dose-Response Modeling (September 1, 2017)
  – Fred Wright (NC State)
    • Overview of the NC State approach to genomic dose-response modeling (September 13, 2017)
  – David Gerhold (NCATS)
    • An Automated Method to Identify Dose-Responsive Genes and Quantitate Points of Departure (PODs) from Transcriptomic Data (September 25, 2017)

• Public comments
Expert Panel

- Carole Yauk (Chair; Health Canada)
- Lyle Burgoon (US Army)
- Ruili Huang (NCATS)
- Kamin Johnson (Dow)
- Rebecca Clewell (Scitovation)
- Jorge Naciff (P&G)
- Setia Pramana (Institute of Statistics)
- James Stevens (Eli Lilly)
- Fred Wright (NC State)

Contributors

- Russell Thomas (US EPA)
- Fred Parham (NIEHS/NTP)
- Deepak Mav (Scione)
- Shyamal Peddada (U. Of Pitt.)

BSC Liaison

- Katrina Waters (PNNL)
• Six sessions
  – Overall approach to genomic dose response studies
  – Study design
  – Filtering of measure features
  – Fitting features to dose response models
  – Determining gene set level potencies
  – Biological interpretation
• Outside speakers were invited to address specific topic areas
• Each session was followed to a 1 hour discussion session
The Speakers

- Russell Thomas (US EPA)
- Fred Wright (NC State)
- Lyle Burgoon (US Army)
- David Gerhold (NCATS)
- Pierre Bushel (NIEHS)
- Jeff Gift (US EPA)
- Keith Shockley (NIEHS)
- Sorin Draghici (Wayne State)
- Woodrow Setzer (US EPA)
- Stephen Edwards (US EPA, now RTI)
- James Stevens (Eli Lilly)
Definition of Terms
**Definition of Terms**

- **Benchmark Response (BMR)**
  - Predefined level of response in a population that is considered to either be adverse or represent significant change

- **Benchmark Dose (BMD)**
  - Model-based, interpolated dose of a test article that produces a level of response defined by the BMR

- **Benchmark dose upper and lower bound (BMD\textsubscript{U}, BMD\textsubscript{L})**
  - Upper and lower bound of the BMD confidence interval that is based on how well the model fit the data

- **Biological effect point of departure (BEPOD)**
  - BMD\textsubscript{L} of the most sensitive "active" gene set
Definition of Terms (2)

- **Feature**
  - The entity that is measured
  - In the case of microarrays it is a probe or probe set that hybridizes to a transcript
  - There can be multiple probe sets for a gene

- **Gene set**
  - Curated group of genes that share an association with specific biological processes (e.g., fatty acid metabolism) or signaling pathways (P53 signaling pathway)
  - Used in combination with gene expression data to provide insight into the molecular level effects that are altered by experimental manipulations
Meeting Sessions

Review of Specific Components of the NTP Proposed Approach
Overall Approach to Genomic Dose Response Studies
Proposed Overall Approach to GDRS

Study Design
Many dose levels, limited biological replication, select target organs/cell types

Filtering Measured Features
Statistical and effect size filter

Fitting Features to Dose-Response Models
Fit multiple models, determine potency using best fit model

Determination of Gene Set Level Potencies
Identify “active” gene sets, report potency for active gene set

Biological Interpretation
Use detailed curation of gene sets to provide toxicological associations/contextualization
Revised Overall Approach to GDRS

**Study Design**
Many dose levels, limited biological replication, select target organs/cell types

**Filtering Measured Features**
Statistical and effect size filter

**Fitting Features to Dose-Response Models**
Fit multiple models, determine potency using best fit model

**Determination of Gene Set Level Potencies**
Identify “active” gene sets, report potency for active gene set

**Biological Interpretation**
Use existing gene sets (e.g., GO) and limit biological interpretation to avoid confusion with hazard identification
Study Design

Filtering Measured Features
Fitting Features to Dose-Response Models
Determination of Gene Set Level Potencies
Biological Interpretation
• Use a “BMD-focused” study design
  – More dose levels fewer biological replicates compare to traditional toxicity studies
  – Example design: 10-12 dose levels, 3 biological replicates/dose group
  – Allows for…
    • Better coverage of the numerous dose-response relationships in each study
    • More confident fits of the data and greater certainty in the BMD estimates for the features
• Use a “BMD-focused” study design
  – More dose levels fewer biological replicates compare to traditional toxicity studies
    • Consider adding additional control animals
  – Example design: 10-12 dose levels, 3 biological replicates/dose group
  – Allows for…
    • Better coverage of the numerous dose-response relationships in each study
    • More confident fits of the data and greater certainty in the BMD estimates for the features
NTP’s Proposed In Vivo Study Design Parameters

- **Sex/Strain/Species:** Male Sprague Dawley Rat
- **Duration:** 5 Days (5 doses, 1 per day, Euthanize 24 hours after last dose)
- **Target Organ Selection:** Liver and expert selected targets based known toxicological effects on structurally related chemicals
- **Top dose selection:** 5 day Maximum Tolerated Dose
Revised In Vivo Study Design Parameters

• **Sex/Strain/Species:** Male Sprague Dawley Rat
  – Perform range finding studies in both sexes and select the sex that is most sensitive

• **Duration:** 5 Days (5 doses, 1 per day, Euthanize 24 hours after last dose)
  – Use pharmacokinetic predictions to determine steady-state timescale for duration determination and time point selection
  – Include an earlier time point (1 day) to allow for temporal analysis that can support causal linkage between late and early effects
    • Start building towards causal linkage to adversity which will be needed for risk assessment

• **Target Organ Selection:** Liver and expert selected targets based known toxicological effects on structurally related chemicals
  – Consider expanding organ collection to allow for future testing

• **Top dose selection:** 5 day Maximum Tolerated Dose
Proposed In Vitro Study Design Parameters

- **Species:** Human
- **Sex:** Determined by availability
- **Duration:** Expert determination
- **Cell Type(s):** Organotypic, Commonly Used, Broad Query Biological Space
- **Top dose selection:** LC20 (where feasible)

No changes proposed by expert panel
Filtering of Measured Features
Proposed Approach to Filtering Measured Features

- One way ANOVA / Fold Change
  - MAQC recommendations for maximizing cross-laboratory reproducibility of genomic data
  - Detects non-monotonic responses
- Specific thresholds will be empirically determined for each technology/platform
- Goal in determining technology/platform-specific thresholds
  - Maximize reproducible signal while minimizing false discovery
Revised Approach to Filtering Measured Features

- **Trend Test / Fold Change**
  - Consider non-parametric trend tests to detect non-monotonic responses

- **Specific thresholds will be empirically determined for each technology/platform**

- **Goal in determining technology/platform-specific thresholds**
  - Maximize reproducible signal while minimizing false discovery
Fitting Features to Dose-Response Models
Proposed Approach to Fitting Features to Dose-Response Curves

- Fit features to 9 parametric continuous models
  - Use models derived from US EPA’s BMDS software
  - Hill, Power, Linear, Polynomial 2, 3, Exponential 2, 3, 4, 5
- BMR = 1.349 x SD of controls
  - Approximates 10% change ~ BMRs used in for dichotomous data (e.g., tumors)
- 2 step process for best model selection
  - Nested Chi Square – Best Poly model
  - Lowest AIC
- From the best fit model determine a BMD, BMD_L and BMD_U
Revised Approach to Fitting Features to Dose-Response Curves

- Fit features to 8 parametric continuous models
  - Use models derived from US EPA’s BMDS software
  - Hill, Power, Linear, Polynomial 2, 3, Exponential 2, 3, 4, 5
  - Add additional polynomial models, but constrain them to allow only one direction change
  - Add a non-parametric modeling option to BMDExpress 2.0

- BMR = 1.349 x SD of controls
  - Consistent with recommendations of US EPA when analyzing continuous data where an adverse effect level is not known

- 1 step process for best model selection
  - Nested Chi-Square — Best Poly-model
  - Lowest AIC (consistent with US EPA's current recommendations)

- From the best fit model determine a BMD, BMD$_L$ and BMD$_U$
  - Consider model averaging
Proposed Approach to Removing Features with Inadequate Model Fits

- Remove features from analysis where best models...
  - Have non-convergent BMD, $\text{BMD}_L$ or $\text{BMD}_U$ values
  - Have a BMD > highest dose
  - Have a nominal global goodness of fit p-value < 0.0001
  - Have a $\text{BMD}_U/\text{BMD}_L > 40$
Final Approach to Removing Features with Inadequate Model Fits

- Remove features from analysis where best models…
  - Have non-convergent BMD, BMD_L or BMD_U values
  - Have a BMD > highest dose
  - Have a nominal global goodness of fit p-value < 0.1

  - Consider using R^2 value instead of or in addition to a global goodness-of-fit p-value
  - Have a BMD_U/BMD_L > 40
Determining Gene Set Level Potencies
Proposed Approach to Identifying “Active” Gene Sets and Determining Potency

• An “active” gene set must...
  – Contain at least 3 genes that correspond to measured features that passed all filters
  – At least 5% populated
  – Have a p-value < 0.05 based on a Fisher Exact Test

• “Active” Gene Set Potency
  – Median gene BMD, BMD$_L$, and BMD$_U$ in a gene set
Revised Approach to Identifying “Active” Gene Sets and Determining Potency

• An “active” gene set must…
  – Contain at least 3 genes that correspond to measured features that passed all filters
  – At least 5% populated
  – Have a p-value < 0.05 based on a Fisher Exact Test
  – Consider using resampling based approaches to determine significant enrichment

• “Active” Gene Set Potency
  – Median gene BMD, BMD$_L$, and BMD$_U$ in a gene set
  – Consider using bootstrapping to determine confidence intervals on gene sets
Study Design
Filtering Measured Features
Fitting Features to Dose-Response Models
Determination of Gene Set Level Potencies

Biological Interpretation
Proposed Approach to Biological Interpretation

- Use expanded version of Hallmark Gene Sets (Liberzon et al., Cell, 2015)
  - Limited redundancy
  - High percentage of genes in each set are regulated at the level of transcript abundance
- Expand the Hallmark Gene Sets
  - Mine the GEO database to identify co-regulated gene sets not currently captured in the Hallmark Gene sets
  - Mine existing phenotypic-anchored signatures such as those that contained in the DrugMatrix database and those from the published literature
  - Remine MSigDB in manner similar to the what was done to create the Hallmark gene sets to identify additional sets that may have been overlooked
- Curate mechanistic and toxicologically relevant interpretations of the “Hallmarks+” gene sets
Revised Approach to Biological Interpretation

- Do not use the proposed approach
- Use existing curated gene set that broadly covers biological space
  - Gene Ontology (GO) Biological Processes, MSigDB C2 gene sets
- Focus on identifying the biologically responsive dose and BEPOD, not hazards
  - Report the lowest 5-10 gene sets, its name and description provided by the curator
  - Do not interpret further
  - Make all data available
- Parallel plan
  - Work within the AOP framework to start associating gene sets with key events
  - Start a formal process for toxicological interpretation GDRS data
• DMPT caused extensive non-neoplastic and neoplastic effects in the liver and respiratory system in a 2-year NTP study performed in F344/N rats
  – **Dose levels:** 0, 6, 20, 60 mg/kg/day by corn oil gavage

• A 5 day genomic dose response study of DMPT was performed in male F344/N
  – **Dose levels:** 0, 1, 6, 20, 60, 120 mg/kg/day by corn oil gavage
  – **Organ for genomic analysis:** Liver
DMPT: Proposed vs Revised Analysis Pipeline

Comparison of Select GO Biological Process BMDs from 5-day GDRS
DMPT: Proposed vs Revised Analysis Pipeline

Apical Potency vs. 5-day GDRS BEPOD

- 1/2 log above and below bile duct fibrosis BMD (2-year study)
- 1/2 log above and below nose, adenoma or carcinoma BMD (2-year study)
- 1/2 log above and below liver, adenoma or carcinoma BMD (2-year study)
- GO Biological Process (proposed pipeline)
  - BMD_L-BMD-BMD_U (5-day GDRS)
- GO Biological Process (revised pipeline)
  - BMD_L-BMD-BMD_U (5-day GDRS)
- MSigDB C2 (proposed pipeline)
  - BMD_L-BMD-BMD_U (5-day GDRS)
- MSigDB C2 (revised pipeline)
  - BMD_L-BMD-BMD_U (5-day GDRS)

With regard to identifying a BEPOD, the overall approach is robust to parameter changes!
BEPODs from 5-day GDRS provide a relatively accurate estimation of the most sensitive apical potency value from guideline toxicity assessments.
Acknowledgements: BMDExpress 2.0

Sciome
Jason Phillips
Dan Svoboda
Arpit Tandon
Deepak Mav
Ruchir Shah
Alex Sedyh
Jason Pirone
Logan Everett

NTP
National Toxicology Program
U.S. Department of Health and Human Services
Fred Parham
Stephen Ferguson
Trey Saddler
Beruk Kiros
Shyamal Peddada
Andy Shapiro
Rick Paules
John Bucher
Ray Tice

Russell Thomas
Jeff Gift
Allen Davis
Louis Olszyk
Josh Harrill

Health Canada
Carole Yauk
Byron Kuo

VISTRONIX
Longlong Yang

ScitoVation
Mel Anderson

Tox Strategies
Julia Rager

Independent Consultant
Bruce Allen
Acknowledgements: Meeting and Webinars

NTP
National Toxicology Program
U.S. Department of Health and Human Services

Mary Wolfe

ICF

Anna Stamatogiannakis
Susan Blaine
Candyn Byrd

NIH
National Institute of Environmental Health Sciences

Predictive Toxicology and Disease Faculty